Direct measurements of meltwater runoff on the Greenland Ice Sheet surface

Laurence C. Smith[1], Kang Yang[1], Lincoln Pitcher[1], Brandon Overstreet[2], Vena Chu[3], Asa Rennermalm[4], Jonathan Ryan[5], Matthew Cooper[1], Colin Gleason[6], Marco Tedesco[7], Jeyavinoth Jeyaratnam[7], Dirk van As[8], Michiel van den Broeke[9], Willem Jan van de Berg[10], Brice Noël[10], Peter Langen[11], Richard Cullather[12], Bin Zhao[13], Michael Willis[14], Alun Hubbard[15], Jason Box[16], Brittany Jenner[17], Alberto Behar[18]

[1]University of California - Los Angeles (UCLA), [2]University of Wyoming, [3]University of California - Santa Barbara (UCSB), [4]Rutgers The State University of New Jersey,, [5]Aberystwyth University, [6]University of Massachusetts, Amherst, [7]The City College of New York, [8]Geological Survey of Denmark and Greenland (GEUS), [9]Utrecht University, [10]Utrecht University,, [11]Danish Meteorological Institute, [12]University of Maryland at College Park, [13]NASA Goddard Space Flight Cente, [14]University of Colorado, [15]Centre for Glaciology, Institute of Geography and Earth Sciences, [16]Byrd Polar Research Center, [17]SonTek, San Diego, [18]NASA Jet Propulsion Laboratory (JPL)

Meltwater runoff from the Greenland Ice Sheet surface influences surface mass balance (SMB), ice dynamics and global sea level rise, but is estimated with climate models and thus difficult to validate. We present a way to measure ice surface runoff directly, from hourly *in situ* supraglacial river discharge measurements and simultaneous high-resolution satellite/drone remote sensing of upstream fluvial catchment area. A first 72-hour trial for a 63.1 km^2 moulin-terminating internally drained catchment (IDC) on Greenland's mid-elevation (1207-1381 m.a.s.l.) ablation zone is compared with melt and runoff simulations from HIRHAM5, MAR3.6.1, RACMO2.3, MERRA-2 and SEB climate/SMB models. Current models cannot reproduce peak discharges or timing of runoff entering moulins, but are improved using synthetic unit hydrograph theory (SUH). Retroactive SUH applications to two older field studies reproduces their findings, signifying that remotely sensed IDC area, shape, and river-length are useful for predicting delays in peak runoff delivery to moulins. Applying SUH to HIRHAM5, MAR3.6.1, RACMO2.3 gridded melt products for 799 surrounding IDCs suggests their terminal moulins receive lower peak discharges, less diurnal variability, and asynchronous runoff timing relative to climate/SMB model output alone. Conversely, large IDCs produce high moulin discharges, even at high elevations where melt rates are low. During this particular field experiment models overestimated runoff by +21 to +58%, linked to overestimated ablation and possible meltwater retention in bare, low-density ice. Direct measurements of ice surface runoff will improve climate/SMB models, and incorporating remotely sensed IDCs will aid coupling of surface mass balance with ice dynamics and subglacial systems.

Ice sheet meltwater runoff | surface mass balance (SMB) | climate models | fluvial catchment | surface water hydrology

Introduction

The production and transport of meltwater (runoff) is an important hydrological process operating on the surface of the Greenland Ice Sheet (GrIS). Total GrIS mass loss from runoff and solid ice dynamics (glacier calving) now exceeds some ~260 Gt/year, contributing >0.7 mm annually to global mean sea level rise (1-3). Since 2009, some two-thirds of this total mass loss has been driven by negative ice sheet surface mass balance (SMB) and associated runoff increases, as calculated from climate/SMB models (3, 4). This runoff passes through supraglacial stream/river networks entering moulins (englacial conduits) and crevasses that connect to the bed (5-9), temporarily influencing basal water pressures and/or ice motion (10-13) and forming a dynamic subglacial drainage system that expels water towards the ice edge and global ocean. The new dominance of runoff as a driver of GrIS total mass loss will likely persist into the future, owing to further increases in surface melting (14), reduced meltwater storage in firn due to

formation of near-surface ice layers (15), and possibly a waning importance of dynamical mass losses as ice sheets retreat from their marine-terminating margins (16). Therefore, the hydrological process of ice surface runoff warrants study, both for basic scientific understanding and to improve representation and/or parameterization of runoff processes in climate/SMB models.

A key uncertainty in climate/SMB projections of future GrIS runoff contributions to global sea level is that estimating runoff requires partitioning of SMB among some poorly constrained processes, with the modeled "runoff" (R) simply an error-sensitive residual of the sum of modeled meltwater production (M), rainfall and condensation, minus modeled retention, refreezing, and sublimation in snow and firn. Representation of these various elements varies by model (see **SI sections 6.1-6.5**) but in all cases R is an error-sensitive residual that is not independently validated with *in situ* field measurements collected on the ice surface. Previous efforts to validate R have used proglacial river discharge (outflow) emerging from the ice edge (7, 17-20), but outflow fundamentally differs from R because it incorporates complex en- and subglacial processes that can delay, remove, or add water, including cavity storage/release, reservoir constrictions, conduit pressurization, subdaily variations in hydraulic potential

---

Significance

Meltwater runoff is an important hydrological process operating on the Greenland Ice Sheet surface that is rarely studied directly. By combining satellite and drone remote sensing with field measurements of discharge in a large supraglacial river, we obtain 72 hours of runoff observations suitable for comparison with climate model predictions. The field observations quantify how a large, fluvial supraglacial catchment attenuates the magnitude and timing of runoff delivered to its terminal moulin and hence the bed. The data are used to calibrate classical fluvial hydrology equations to improve meltwater runoff models, and demonstrate that broad-scale surface water drainage patterns that form on the ice sheet powerfully alter the timing, magnitude, and locations of meltwater penetrating into the ice sheet.

**Reserved for Publication Footnotes**
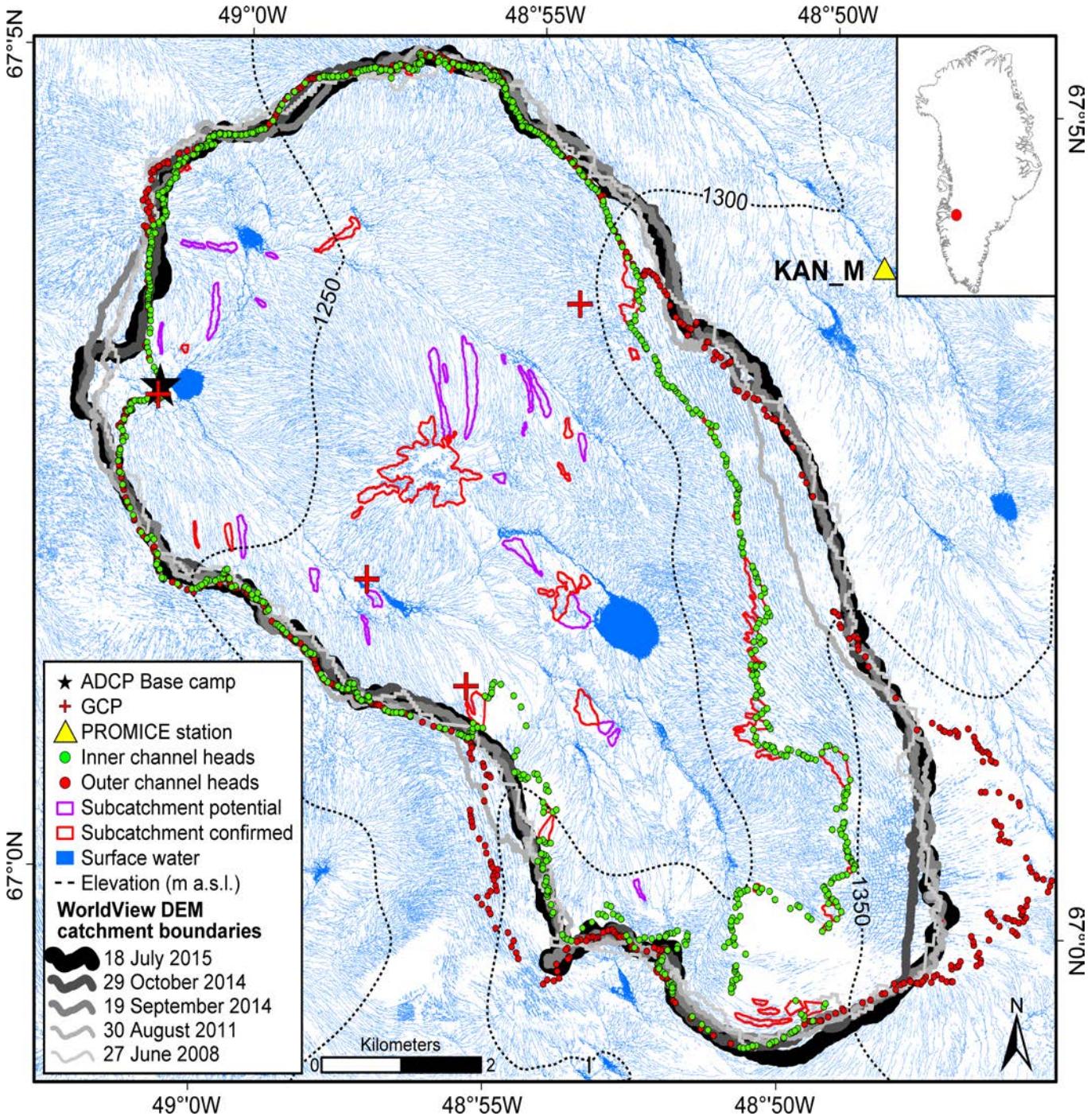
Fig. 1. WorldView-1/2 satellite mapping of Rio Behar catchment, a moderately sized (63.1 km²) internally drained catchment (IDC) centrally located in a melt-intensive area of the Greenland Ice Sheet (inset). From 20-23 July 2015, we collected 72 hours of continuous in situ ADCP discharge measurements in the main-stem supraglacial river (Rio Behar) ~300 m upstream of the catchment's terminal moulin at our base camp (black star, 67.049346N, 49.025809W). Measurements of ice surface ablation were collected at base camp (inside the black star) and by the PROMICE KAN_M automated weather station (yellow triangle). GPS-surveyed red tarpaulins were used as ground control points (GCP, red plus symbols) to aid satellite and UAV image geolocation and georectification. Eight years of topographic catchment boundaries delineated from WorldView satellite stereo-photogrammetric digital elevation models (DEMs, multi-shaded gray lines) establish overall catchment stability from 2008-2015; the 18 July 2015 DEM boundary, adjusted for small areas of stream piracy, was used for calculations presented in this study (thick black line, 63.1 km²). Manually identified stream channel heads (headwater channel incision points) mapped in the 18 July 2015 satellite image constrain minimum (green circles – inner) and maximum (red circles - outer) plausible catchment boundaries, respectively. The minimum boundary eliminates crevasse fields in the southeast catchment headwater area. Polygons bound confirmed (red polygon) and potential (purple polygon) internally drained subareas (i.e. small internal moulins) not draining to the large terminal moulin. Four small, non-draining supraglacial lakes were fully integrated into the stream/river network with no impoundment of flow. This map was created using DigitalGlobe, Inc. imagery.

gradient, basal melting, and subglacial aquifers (5, 10, 17, 21-23). Furthermore, basin delineations for proglacial river outlets have

high uncertainty (7, 17, 24), are keenly sensitive to user choice of a hydraulic potential parameter (i.e. the *k*-value (25, 26)),
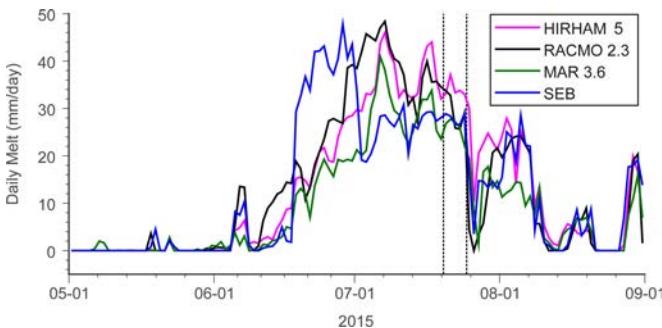
Fig. 2. The 20-23 July 2015 field experiment (dashed lines) was timed for late July near the end of the peak runoff season, when Rio Behar catchment was bare ice, the seasonal surface drainage pattern was fully developed, and prior to the onset of cooler temperatures and reduced melting in August. Colored lines show daily melt rates for the HIRHAM5, RACMO2.3, MAR3.6, and Point SEB climate/SMB models, melt rate is not supplied by MERRA-2.

and are vulnerable to water piracy between adjacent basins (27, 28). Proglacial river discharge measurements can also suffer large uncertainty due to heavy sediment loads, braided channels, and mobile beds (29). In short, proglacial outflow does not confidently reflect the timing of SMB and runoff processes operating on the GrIS surface, especially at diurnal time scales.

At the present time, climate/SMB models contain little or no provision for retention and/or refreezing of runoff in bare ice (i.e. either on or below the ice surface), or for flow routing (lateral transport) of runoff over the ice surface to moulins. Instead, residual M converts instantly to R and is assumed to depart the ice surface. This is acceptable for estimating net SMB but not for estimating the timing and volume of runoff delivered to moulins, the dominant pathway linking supraglacial with subglacial hydrologic systems (7, 21, 24). This in turn clouds understanding of the interplay between SMB and ice dynamics, especially at short time scales. Moulins inject surface runoff into a transient, subglacial hydrologic system exerting primary control on diurnal to multi-day changes in ice sheet basal motion and water pressure (11, 12, 30-33). Sub-daily delays or lags between the timing of surface melt and basal water pressures are often used to infer capacity of the subglacial drainage system, yet supraglacial routing delays receive little or simplified treatment (9, 10, 31, 34).

Finally, solar radiation supplies most energy for melting ice on the GrIS margins and bare-ice ablation zone, followed by the turbulent flux of sensible heat (35, 36). As a result, temporal scales governing energy and mass exchange between the atmosphere and ice surface range from seconds (for turbulent eddies) to daily and monthly for net radiative surface energy balance. Because solar radiation dominates melting, it is imperative to resolve the effect of diurnal cycles in the surface energy balance on surface runoff patterns. The diurnal time scale is especially important for runoff generation in the mid-elevation ablation zone, where daytime melting is interrupted by nighttime freezing (37), causing heat loss from the ice surface and potential refreezing of melt-water. Diurnal variations in runoff also influence ice dynamics, because ice motion accelerations are driven by variability in melt-water input (10, 12). Meltwater alternatively flows from subglacial channels into the distributed basal system during intervals of high supply/high pressure, and from the distributed system into chan-nels during intervals of low supply/low pressure (33, 38, 39). This diurnal pressurization of the distributed system drives diurnal variations in ice velocity. Numerical modeling shows increases in diurnal ice motion and a slight increase in annual mean velocity when diurnal variations in surface runoff input are considered (40).

In sum, climate/SMB models are essential tools for sim-ulating SMB runoff inputs to subglacial systems and to the

global ocean (41, 42), but currently lack validating field mea-surements of runoff timing and quantity, especially over short time scales. To address these challenges, we present a field-based approach to measure R directly on the ice sheet surface – prior to en- and subglacial interferences – at the scale of an internally drained supraglacial catchment (IDC). IDCs are defined by flu-vial supraglacial stream/river networks, which dominate surface drainage patterns of the southwestern GrIS (43). They have areas of order $\sim 10^1$ -10^2 km^2, a geographic scale comparable to the grid cells of most regional climate/SMB models. The field procedure is demonstrated for a representative IDC having an area of 63.1 km^2 (our best estimate of catchment area, with upper and lower uncertainty bounds of 69.1 km^2 and 51.4 km^2, respectively), hereafter called Rio Behar catchment in honor of the late Dr. Alberto E. Behar[†] (**Figure 1**). Spanning a 1207 – 1381 m elevation range, Rio Behar catchment is located just below the long-term equilibrium line ($\sim$1500 m a.s.l. in this area (34)), experiences seasonal melting from June through August of each year, and is centrally located in one of the highest runoff-producing regions of the GrIS (3, 14). Our field trial was conducted in late July 2015, near the end of the peak runoff season when the region's supraglacial stream/river networks are fully developed, yet prior to the onset of reduced melting in August (**Figure 2**).

Conceptually, our approach is simple, requiring only hourly measurements of discharge in an IDC main-stem supraglacial river (i.e. to measure the volume of runoff physically departing the source catchment) and high-resolution mapping of the IDC's con-tributing upstream catchment area. Note that "runoff" has units of depth per model time step in gridded climate model output (L T^{-1}, typically mm d^{-1} or mm hr^{-1}) but units of discharge when obtained from in situ measurements (L^3 T^{-1}, typically m^3 s^{-1}). Remotely sensed catchment area (L^2, typically km^2) is required for conversion between the two units of runoff.

We measured discharge hourly in the main-stem supraglacial river of Rio Behar catchment for 72 hours from 20 – 23 July 2015, by deploying an RTK GPS SonTek RiverSurveyor® Acous-tic Doppler Current Profiler (ADCP) from a bank operated cableway suspended across the river immediately upstream of its descent to the catchment's terminal moulin (**Figure S1**). During the same period, we obtained high-resolution images from the WorldView-1 and WorldView-2 satellites (resolution 0.5 m panchromatic, 2.0 m multi-spectral) and a custom-made fixed-wing drone (Unmanned Aerial Vehicle or UAV, visible band, resolution 0.3 m). These acquired images were used to map contributing Rio Behar catchment boundaries, surface drainage pattern, and snow cover. Topographic divides of Rio Behar catchment were delineated from a high-resolution digi-tal elevation model (DEM) of the ice surface, derived stereo-photogrammetrically from a WorldView-1 image pair acquired 18 July 2015. The long-term stability of this divide was established from older WorldView image pairs beginning in 2008 (**Figure 1**). The 2015 topographic boundary was later manually adjusted for small areas lost (2.7 km^2) or gained (0.8 km^2) due to stream piracy (breaching) across divides, and for small internal sub-areas drain-ing to internal moulins (1.6 km^2). Intersection of this corrected catchment area (63.1 km^2), and also its maximum plausible extent (69.1 km^2, identified by mapping outer channel heads, **Figure 1**) and minimum plausible extent (51.4 km^2, identified by mapping inner channel heads and removal of 4.1 km^2 of crevasse fields, see **Figure 1** and **SI section 3.2**) with gridded outputs from the HIRHAM5, MAR3.6.1, RACMO2.3, MERRA-2 and Point SEB climate/SMB models enables a first direct comparison between modeled and measured on-ice R for Rio Behar catchment (**Figure 3; Figure 7a**).

Fig. 3. Hourly supraglacial runoff R from Rio Behar catchment obtained from in situ ADCP discharge measurements (red) and as estimated by five climate/SMB models (colored envelopes) during the 20-23 July 2015 field experiment. Observed runoff is attenuated and delayed relative to modeled runoff due to non-representation of fluvial transport (routing) in current models. An exception is MAR3.6 (green) which uses a simple delay-to-ice-edge assumption thus greatly smoothing the diurnal runoff signal. Units of R in climate/SMB models (mm hr^{-1}) are converted to discharge (m^3 s^{-1}) by multiplication with remotely sensed catchment area (Figure 1), enabling direct comparison with ADCP measurements. The uncertainty bounds shown for modeled R thus reflect catchment area uncertainty, with centerlines denoting the optimal area estimate of 63.1 km^2 and upper and lower uncertainty reflecting the maximum and minimum plausible area estimates of 69.1 km^2 and 51.4 km^2, respectively. Error bars for in situ data are standard deviations calculated from multiple ADCP profiles collected within each measurement hour.

Results

Comparison of our hourly discharge measurements (**Table S1**) with hourly climate/SMB model outputs of catchment R quantifies the attenuation and delay of observed R delivered to the Rio Behar catchment terminal moulin (**Figure 3**). Because evacuation of runoff requires physical passage through the IDC's fluvial drainage pattern, some duration of time must pass between the timing of peak R generated across the IDC and the timing of peak R (discharge) received by the moulin. This duration is called "time-to-peak" (t_p, in hours) in traditional terrestrial hydrograph analysis (44, 45). In general, time-to-peak delays will increase for catchments having larger and/or more elongate areas and lower stream densities, with soil porosity, topographic slope and land cover contributing factors (45). For a given uniform depth of R generated across the catchment, larger catchments produce greater total discharge and peak discharge (Q_{pk}) than smaller catchments, due to their larger source areas. Applied to southwest Greenland, where most IDCs have areas of tens of square kilometers (43), these fluvial geomorphology processes are thus intrinsic to the scale of a climate/SMB model grid cell.

To demonstrate how influential fluvial supraglacial catchments are to the timing (t_p) and peak discharge (Q_{pk}) of GrIS meltwater runoff delivery to moulins, we use our Rio Behar discharge measurements to calibrate a simple lumped (catchment-scale) morphometric routing model for use on the ablating ice surface, the synthetic unit hydrograph or SUH (see **SI Methods 4**). Three advantages of the SUH routing model are that it isolates the impact of basic IDC properties (area, stream length, and shape) on t_p and Q_{pk} delivered to the catchment outlet (here, the terminal moulin) which can be obtained with remote sensing; it does not require use of digital elevation models (which are acutely sensitive to choice of a depression-filling threshold, and don't always reflect true surface drainage patterns (46)); and it is designed to be transferable to ungauged catchments.

Extension of our field-calibrated SUH to a broad-scale (13,563 km^2) remotely-sensed map of 799 surrounding IDCs (43)

quantifies temporal and spatial heterogeneities in runoff delivery to terminal moulins due solely to differences in IDC areas, river-lengths and shapes (**Figure 4**). For a theoretical unit runoff depth of 1 cm (i.e. a 1 cm layer of water assumed to materialize uniformly across the ice sheet surface in 1 hour), catchment-induced time-to-peak delays would range from as low as 0.4 h to as high as 9.5 h, due solely to varying IDC areas, stream lengths, and shapes (**Figure 4a**). Peak discharges entering moulins would range from as low as 0.7 m^3 s^{-1} to as high as 53.0 m^3 s^{-1} (**Figure 4b**), again due solely to these basic fluvial catchment properties that are not currently represented in climate/SMB models.

A more realistic scenario, using climate/SMB model outputs of melt production M and a Gamma function to synthesize each IDC's unique SUH (47) (**SI Methods 5**) yields similarly heterogeneous spatial patterns not present in gridded climate model output (**Figure 5, Figure S11**). These heterogeneities include large discharges (>20 m^3 s^{-1}) entering moulins at high elevations on the ice sheet (>1500 m a.s.l.) despite low melt rates there, due to presence of large IDCs (43, 48). Importantly, peak moulin discharges are significantly reduced if climate/SMB model output is subjected to unit hydrograph theory (**Figure 5c**), rather than the practice of instantaneously aggregating model output within each IDC (33, 49) (**Figure 5b**). The opposite is true at night, when modeled melt and instantaneous area-aggregated runoff shut down but SUH-routed runoff is high (**Figure S11**). Averaging across all 799 IDCs (including all small catchments) mean Q_{pk} is reduced by 13.5 + 10.0% when climate/SMB model output is subjected to SUH routing (**Figure 6**). Diurnal variability in Q_{pk} is reduced by 15.1 +12.5%, and the mean timing delay between peak melt production and peak moulin discharge lengthens by 2.9 + 2.8 hours. For the larger IDCs (>30 km^2, n=122), for which routing delays are greatest, these averages rise to 30.4 + 9.1%, 37.0 + 12.0%, and 5.1 + 4.6 hours, respectively.

While these numbers should be viewed cautiously as our SUH model depends in part on parameters calibrated only at Rio Behar catchment, a successful retroactive application of SUH to
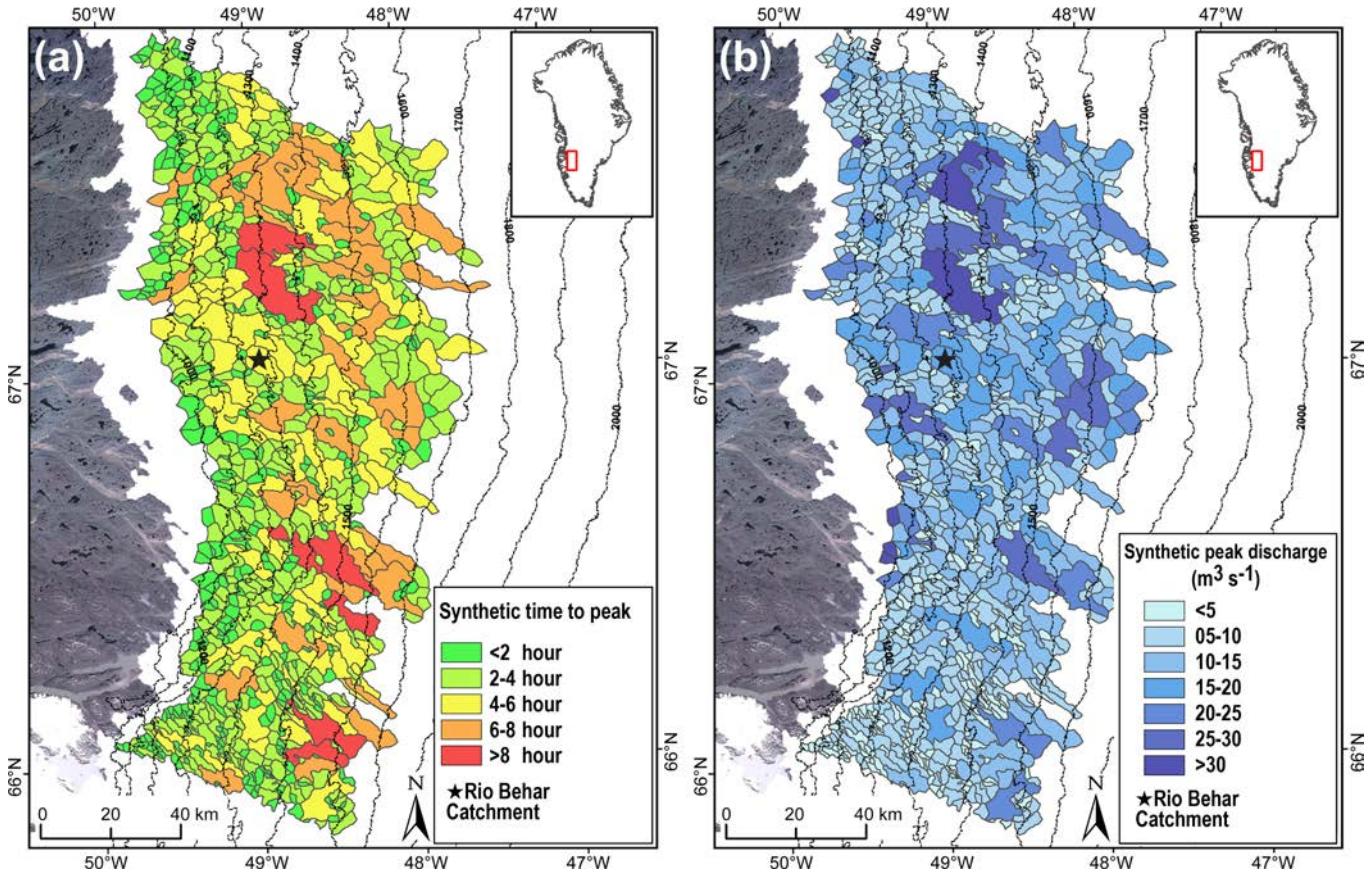
Fig. 4. Application of our field-calibrated synthetic unit hydrograph (SUH) routing model to 799 remotely-sensed IDCs on the southwest GrIS (gray polygons, mapped previously from a 19 August 2013 panchromatic Landsat-8 image (43)) illustrate how fluvial, supraglacial internally drained catchments (IDCs) impart spatially heterogeneous modifications to meltwater runoff delivered to terminal moulins and hence the bed. Each IDC contains a remotely sensed supraglacial river network (not shown for visual clarity) terminating in a major, catchment-terminating moulin. These theoretical SUH maps assume a spatially uniform 1 cm deep layer of meltwater released over a duration of 1 hour, and isolate the influence of IDC area, shape, and stream length on (a) time-to-peak delays of peak runoff arrival at each catchment's terminal moulin (t_p, in hours); and (b) magnitude of peak discharge received at each catchment's terminal moulin (Q_{pk}, m³ s⁻¹). More realistic maps, forced by climate/SMB models, appear in Figure 5 and Figure S11.

the IDCs of two older field studies (8, 32) is encouraging (**SI Section 5.3**). Depending on choice of input climate/SMB model, SUH-estimated peak runoff times for a 1.1 km² IDC nearly 300 km distant from Rio Behar catchment range from 16:00 to 20:00 (local Greenland time), comparable to 16:30-17:00 observed in field observations acquired in August 2009 (8, 32) (**Table S4**). For an 18.2 km² IDC approximately 14 km distant from Rio Behar, SUH-estimated peak runoff times range from 17:00-22:00, comparable to field measurements of 18:00-20:00 acquired in late July/early July 2011 (32). Such independent reproductions of runoff timing delays measured at other times and locations on the ice sheet suggest utility of SUH elsewhere on the southwest GrIS ablation zone. However, collection of additional supraglacial discharge datasets, especially from large IDCs and colder regions, are needed for further calibration and validation of the SUH approach.

With regard to the absolute magnitudes of measured versus modeled runoff, comparison of our cumulative ADCP discharge measurements with cumulative modeled R over our 72-hour field experiment finds that climate/SMB models overestimated R by +21% to +58% for this particular location and time on the ice sheet (for a 5-model average, assuming lower and upper constraints on watershed extent, respectively). Taken separately, 4/5 models overestimated R (**Figure 7a**). Similarly, 4/4 models (for which melt M is available) overestimated ice surface lowering (ablation), if their outputs of M are compared with in situ

ice surfacing lowering measurements collected from 15 ablation stakes at our base camp, and sonic surface lowering data from the nearby PROMICE KAN_M automated weather station (AWS) (**Figure 7b, Table S5**). This conclusion holds regardless of whether the density of solid ice (0.918 g cm⁻³) is used to convert M to units of ice thickness equivalent, or a lower, near-surface ice density (0.688 g cm⁻³) averaged from ten shallow cores drilled at our base camp (50). Point-based ablation measurements have known limitations (51), but both field datasets display less ice surface lowering than modeled M (**Figure 7b**), similarly to how the models overestimate R (**Figure 7a**).

One interpretation of **Figure 7** is that the models overestimated M, and hence R. However, examination of modeled vs. in situ AWS surface energy balance (**SI section 6.8**) reveals that modeled energy balance components closely matched in situ AWS measurements. In general, RACMO2.3 albedo, radiation, and turbulent fluxes track AWS observations too well to advance model overestimation of surface energy receipt as the leading explanation for model overestimations of ice ablation and R (**Figure S9**). For example, the radiative effects of clouds (52) may have contributed slightly to model overestimation of R during the third day of the field experiment, but not the first two days when the sky was clear (**Figure S9**). Importantly, the Point SEB model is driven purely by AWS measurements yet similarly overestimates observed surface ablation and R like other, reanalysis-driven models (**Figure 7b, 7a**).
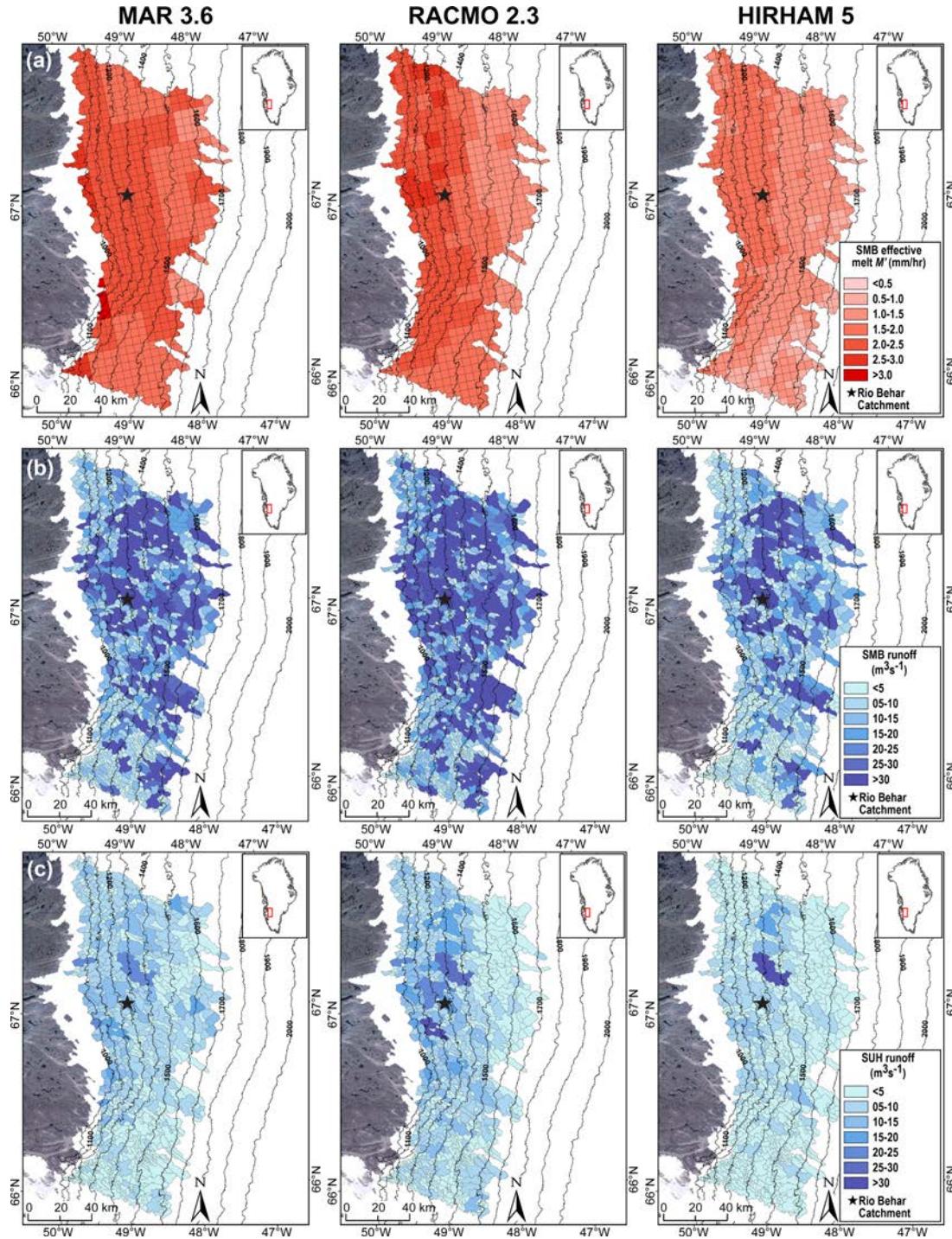
Fig. 5. Supraglacial internally drained catchments (IDCs) modify the timing and magnitude of runoff delivered to terminal moulins, as demonstrated here at 1400 local western Greenland time on 21 July 2015, using (a) MAR3.6, RACMO2.3, and HIRHAM5 climate/SMB model outputs of corrected meltwater production (M', see SI section 4.3) to estimate: (b) instantaneous area-integrated runoff; and (c) more realistic, SUH-routed runoff. MERRA-2 is not shown because it does supply M; Point SEB is not shown because its output is not gridded. The boundaries of 799 IDCs (gray polygons) were mapped previously from a 19 August 2013 panchromatic Landsat-8 image (43). Each IDC contains a remotely sensed, moulin-terminating supraglacial river network (not shown for visual clarity). Climate/SMB model output M' has units of water depth equivalent (mm hr^{-1}), which converts to runoff in discharge units (m^3 s^{-1}) following intersection with IDC catchment boundaries (b), (c). Black star at approximately 67N, 49W denotes the Rio Behar IDC. In both (b) and (c) large IDCs enable large moulin discharges above 1500 m a.s.l. elevation, despite lower overall melt rates. SUH routing (c) yields lower peak moulin discharges at this time of day than instantaneous area-integrated runoff (b), because SUH requires more time for runoff to travel through fluvial supraglacial stream/river networks. A companion nighttime version of this figure ten hours later (24:00, see Figure S11) shows the opposite effect, with shutdowns in (a) and (b) but high moulin discharges in (c).

All of this suggests some meltwater loss or retention process that is external to the "skin" surface energy balance allocated at the top of the ice surface. We hypothesize that subsurface melting (53) and subsequent retention and/or refreezing of meltwater in
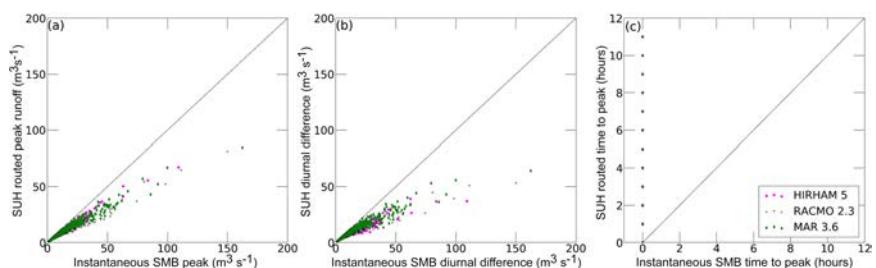
Fig. 6. Comparison of SUH-routed runoff (Figure 5c) versus instantaneous area-integrated runoff (Figure 5b) for all 799 IDCs: (a) peak moulin discharge; (b) diurnal difference between maximum and minimum moulin discharge; and (c) time delay between peak melt production across the catchment and peak discharge at the terminal moulin. Applying SUH routing to climate/SMB model output yields lower peak discharges; suppressed diurnal variability; and delayed, asynchronous timing of peak runoff delivered to large catchment-terminating moulins.
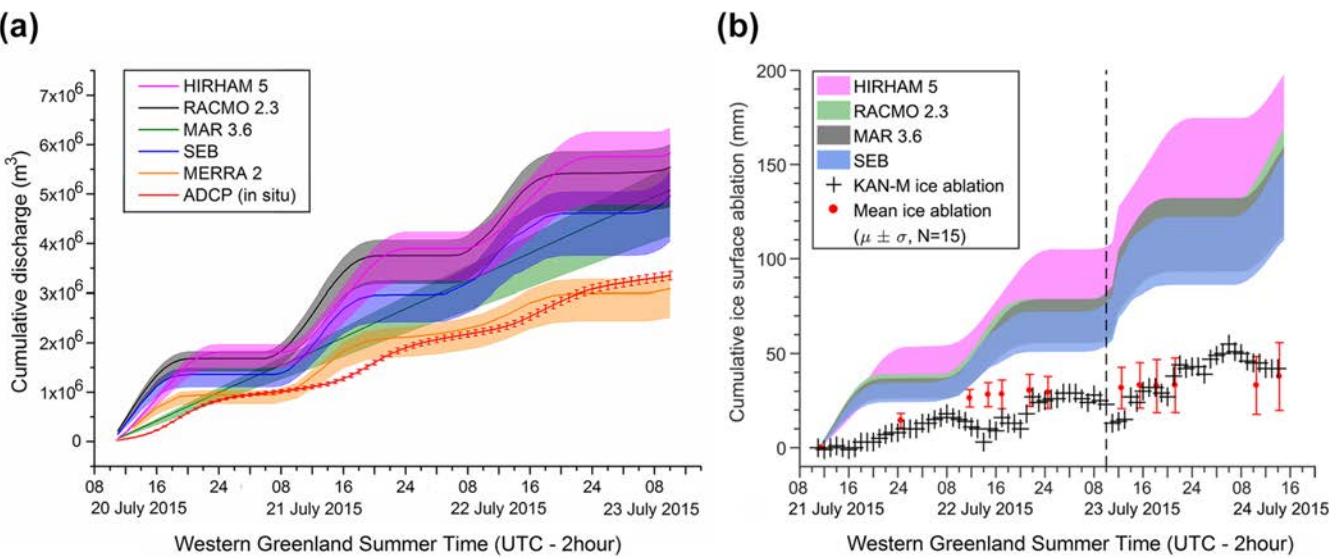

Fig. 7. : Climate/SMB model simulations of modeled vs. measured (a) runoff; (b) ice surface lowering (ablation) during the 20-23 July 2015 field experiment. (a) Cumulative hourly supraglacial runoff R from Rio Behar catchment as measured from in situ Acoustic Doppler Current Profiler measurements (ADCP, in red) and as estimated by five climate/SMB models (colored envelopes). Note that values of cumulative modeled R (m^3) derive from summation of hourly discharges ($m^3 s^{-1}$), which are obtained by multiplying climate/SMB model outputs with the remotely-sensed catchment area(s) of Figure 1. Upper and lower uncertainty bounds in modeled R thus reflect catchment area uncertainty, with centerlines denoting the optimal area estimate of 63.1 km^2 and upper and lower uncertainty bounds reflecting maximum and minimum plausible catchment area estimates of 69.1 km^2 and 51.4 km^2, respectively. Error bars (red) for our in situ data denote: (a) Cumulative standard deviations calculated from multiple ADCP supraglacial river discharge measurements collected within each measurement hour; and (b) cumulative ice surface lowering measurements as measured manually at fifteen ablation stakes in Rio Behar base camp (mean values also shown). Upper and lower uncertainty bounds in modeled ice ablation reflect assumptions either solid ice (0.918 g cm^{-3}) or lower observed (0.688 g cm^{-3}) (50) bare ice density to convert model outputs of M from units of liquid water equivalent to solid ice equivalent. Vertical dashed line in (b) indicates time of cessation of ADCP discharge experiment in (a). MERRA-2 is not shown because M is not supplied by MERRA-2.

porous, low density bare ice (called "weathering crust" (6, 50, 54)) may contribute to or explain the observed discrepancies between modeled M and R, and measured ice surface lowering and supraglacial river discharge, respectively. Runoff infiltration into crevasses (8) cannot explain the observed runoff deficit, as crevassed areas are eliminated from our minimum bounding catchment map (51.4 km^2) and are thus already included in the lower model uncertainty bounds of **Figure 7a** (and **Figure 3**). While the possibility of additional, missed leakage cannot be fully ruled out, there is no evidence for this in our high-resolution UAV imagery (**Supplemental Discussion I, Figure S3**). Missed meltwater retention in seasonal snow also seems unlikely: the climate/SMB models indicate bare ice, and snow classifications from our UAV mapping and two WorldView-2 images confirm that Rio Behar catchment had less than < 6.5% snow cover at the time of our field experiment, and perhaps as little as 0.9% (**SI section 3.4, Figure S4**). Remotely sensed retrievals of lake volume storage rule out the possibility of runoff impoundment in four supraglacial lakes contained within Rio Behar catchment (**SI section 3.3**). The remaining hypothesis, i.e. of water

retention/refreezing in the bare-ice weathering crust, is explored further in the Discussion and in **SI**.

Regardless of mechanism, a first-order, empirical correction for any missed retention processes and/or model overestimations of M for Rio Behar catchment during our field experiment is supplied by a set of empirical, model-specific runoff coefficients relating observed runoff R to modeled melt production M for HIRHAM5, MAR3.6.1, RACMO2.3 and SEB (**Table S3**). No values are supplied for MERRA-2 as M is not an output of this model. While these coefficients are computationally identical to how runoff coefficients are calculated for terrestrial catchments (i.e. river discharge divided by catchment water input), they also include any missed model over/underestimation of M and are more properly treated as correction factors for climate/SMB models instead of traditional runoff coefficients. For bare-ice surface conditions similar to those observed at Rio Behar catchment during our field experiment, these correction factors may be multiplied by M to obtain alternate, lower estimates of M (here termed effective melt M') in addition to standard model output.

Discussion and conclusions

While the field protocol presented here is currently logistically impractical for sustained monitoring or deployment at numerous sites, it offers a useful, and perhaps only direct way to independently measure supraglacial R for validating climate/SMB models used to simulate ice sheet runoff and associated inputs to subglacial and marine systems. Our provision of field-calibrated, model-specific runoff coefficients and SUH parameters offers an initial step in this direction, enabling generation of synthetic unit hydrographs, peak moulin discharges (Q_{pk}), and runoff time-to-peak delays to moulins (t_p), from standard climate/SMB model outputs of melt production M (**Table S3**) at a time of peak drainage efficiency on the ice sheet surface.

Our retroactive testing of SUH runoff timing delays against in situ observations of two earlier field studies (8, 32) conducted in different years, locations, and elevations than Rio Behar catchment suggests plausible transferability of SUH to other areas of the GrIS ablation zone. One reason for this success may be that only three SUH parameters (C_p, C_t and m) require in situ calibration – the others (t_p and h_p) derive purely from remotely sensed catchment characteristics and are thus recalibrated individually for each IDC. That said, further field experiments are needed at other locations and times on the ice sheet to derive additional runoff coefficients and SUH parameters for differing surface conditions. Hydrological measurements from Haut Glacier d'Arolla, Switzerland, for example, suggest that earlier in the runoff season the presence of snow also suppresses diurnal contrasts and introduces delays between peak melt production and peak moulin discharge (55). Similarly, the seasonal evolution of supraglacial stream/river drainage networks may influence early-season runoff coefficients and the values of C_p, C_t and m presented here, due to lower stream density and/or temporary retention of runoff in slush and seasonal snow (43). Note that the most likely outcome of these processes would be to further delay runoff delivery to moulins (**Figure 6c**), further suppress diurnal variability (**Figure 6b**) and further suppress peak moulin discharges (**Figure 6a**), rendering conservative our scientific conclusions about the influence of fluvial supraglacial catchments on meltwater delivery to moulins and the bed.

The field measurements and SUH calculations presented here illustrate the critical importance of IDCs in modulating the timing and magnitude of runoff evacuated off the ice surface to moulins (**Figure 4, Figure 5, Figure S11**). Previous studies have shown the importance of filling and draining supraglacial lake basins (20, 30, 56), but even in the absence of lake basins, runoff becomes unevenly redistributed over space and time due to water capture and transport through fluvial supraglacial stream/river networks. Based on our observed values of C_p and C_t these catchment-scale processes on ice are not unlike those on land (**SI section 5.1**), despite known hydraulic differences between supraglacial and terrestrial channels (57).

Because IDC areas vary greatly and moulins convey meltwater quickly to the bed (32), the timing and volume of runoff received at the bed is thus arrhythmic in time and heterogeneous in space, unlike outputs from gridded climate/SMB models (**Figure 5, Figure S11**). Sub-daily time lags between surface climatology and melt-induced ice motion are established first on the ice surface, that might otherwise be attributed to en-and/or subglacial delays or modes in melt-induced ice motion (8, 11, 12, 25, 30, 56), basal pressure (23), or subglacial drainage capacity (34). Diurnal variability in moulin discharge is lower than that of climate/SMB modeled runoff fields, potentially reducing rates of inferred subglacial channelization (40). Where the diurnal variability of meltwater delivered to the bed is dampened by surface routing delays, there should be an impact on ice sliding velocities, especially at higher elevations on the ice sheet. Using climate/SMB runoff to drive ice dynamics models in such areas could thus overestimate diurnal subglacial pressure variability,

leading to small overestimations in the diurnal range of ice velocities and perhaps annual mean velocity as well. Conversely, large IDCs have capacity to amplify moulin discharge, including at high elevations where melt rates are low but IDCs are large (43, 48), especially if moulins are first initiated through hydrofracturing and drainage of interior-advancing supraglacial lakes (21, 24, 58) and subjected to extreme and/or sustained melt events (59). In sum, the supraglacial drainage pattern of the GrIS surface influences a host of important subglacial processes, especially at short time scales.

Our finding that modeled and observed surface energy balance largely agree (**Figure S9**) yet both overestimate observed ice surface lowering and runoff (**Figure 7**), leads us to hypothesize that subsurface melting and delay/retention/refreezing of meltwater in porous, low density weathering crust may be an important bare-ice physical process not represented in the climate/SMB model simulations presented here. Shortwave radiation penetration and subsurface melting of bare ice certainly promotes development of weathering crust (6, 53, 54) at our study site (**Figure S10**), which is characterized by abundant cryoconite holes and porous, water-saturated, low density bare ice at least 1.1 m deep (50). Ablating weathering crust typically experiences less surface lowering than expected from skin surface energy balance calculations alone, owing to internal melt within the subsurface ice matrix (50, 60, 61). Any meltwater retained within this porous medium, for example due to deepening of the crust, enlargement of cryoconite holes, or enlargement of pore space volume, would result in model overestimation of R because current modeling schemes do not permit water retention in bare ice. Moreover, any refreezing of this meltwater (which we observed nightly during the field experiment) requires that it re-melt to become true runoff, consuming additional melt energy not currently allocated in energy balance models for the bare-ice zone. Any model that correctly quantifies surface melt energy but does not simulate these processes will overestimate both ice surf---ace lowering and runoff (**SI Supplementary Discussion I**).

While mismatched scale and timing preclude direct comparison of our field results with GRACE (Gravity Recovery and Climate Experiment) satellite gravity data, we note in **SI** that two previously published, sector-aggregated GRACE observations similarly show less actual mass loss than simulated by climate/SMB models (SMB-D) in some key melt-intensive sectors, including ours in southwest Greenland (62, 63) (**SI Supplementary Discussion II**). However, we are reluctant to draw general conclusions about climate/SMB model performance at other times or locations on the GrIS, owing to the short duration and small geographic area (relative to model domains) of our field experiment. The observed spread in modeled runoff estimates for Rio Behar catchment (**Figure 7a**) is consistent with a broader inter-comparison of modeled outputs across the GrIS, including heightened model uncertainty in the ablation zone (64). New field experiments are needed to determine how to refine climate/SMB model simulations of ice surface ablation and runoff in the bare-ice zone, as well as remote sensing SMB estimates that use satellite/airborne altimetry measurements of ice surface lowering.

Regardless of absolute magnitudes of R, the timing and amplitude of meltwater runoff is clearly modified by fluvial catchment processes operating on the GrIS surface. Lateral flow routing through internally drained catchments predictably delays the arrival, reduces the peak discharge, and suppresses the diurnal variability of R entering moulins. Large catchments yield high moulin discharges, even at high elevations where overall melt rates are low. These realities, together with possible delays/retention/refreezing of runoff in the bare-ice ablation zone weathering crust, signify that supraglacial drainage processes critically preconfigure the timing and flux of meltwater delivered to the bed. Incorporating fluvial catchments, hydrologic theory, and

Supporting Information (SI)

(for Smith et al., "Direct measurements of meltwater runoff on the Greenland ice sheet surface")

Table of Contents

* * *

Methods 1. Field measurements and data processing

Rio Behar catchment is a moderate sized supraglacial IDC (internally drained catchment) located on the southwestern GrIS surface (**Figure 1**; see also small black star, **Figure 4, Figure 5, Figure S11**). Owing to rapidly melting conditions on the ablation zone, permanent gauging installations are infeasible. From July 17-24, 2015 we established a temporary base camp and supraglacial river discharge gauging installation in the main-stem supraglacial river (termed "Rio Behar") at 67.049598N, -49.0201453W, immediately downstream of the confluence of Rio Behar catchment's two largest tributaries and lake spillway, approximately 300 m upstream of the catchment's terminal moulin.

78   1.1. Acoustic Doppler Current Profiler (ADCP) field experiment:

79   For 72 continuous hours on July 20-23, 2015 we collected in situ Acoustic Doppler Current

80   Profiler (ADCP) measurements of supraglacial river discharge (Q) in Rio Behar, using a bank-

81   operated cableway system anchored to the ice surface (**Figure S1**). The cableway was used to

82   repeatedly tow a SonTek® M9 ADCP mounted on a Hydroboard II back and forth across the

83   river channel, thus obtaining hydrographic profiles of changing channel cross-section, wetted

84   perimeter, and velocity 3-6 times per hour beginning 1:16:34 PM (UTC) July 20, 2015 and

85   ending 12:12:11 PM (UTC) July 23, 2015 (local western Greenland time was UTC-2:00). ADCP

86   measurements were acquired hourly to capture a wide range of discharge values and to avoid

87   reliance on a stage-discharge rating curve (from field observations in 2015 and previously,

88   ongoing thermal erosion of large GrIS supraglacial river channels quickly renders rating curves

89   obsolete). This requirement of around-the-clock in situ ADCP measurements, rather than

90   simply a water level recorder, is the reason for the 72-hour duration of the field experiment.

91
92   **Figure S1. Measurements of supraglacial river discharge were collected continuously for 72 hours**

93   **approximately 300 m upstream of the Rio Behar catchment terminal moulin, using an Acoustic**

94   **Doppler Current Profiler (ADCP) operated by rotating shifts of technicians safely tethered to the ice. A**

95   **bank operated cableway system was used to repeatedly tow the ADCP back and forth across the**

96   **channel (photo by Åsa Rennermalm).**

97   1.2 Acoustic Doppler Current Profiler (ADCP) data collection and processing:

98   Given that the ADCP discharges represent the core dataset of this study, the measurements are

99   described in detail. The Sontek® M9 ADCP uses real-time-kinematic (RTK) GPS precision and

100   Doppler technology to measure channel cross sectional area and velocity. The Doppler principle

101 assumes that suspended particles in the water column travel at the same speed as water and
102 the change in frequency detected by each acoustic transducer is translated into a velocity using
103 the speed of sound, calculated from temperature and salinity.

104 While in transect, the ADCP determines its position using (1) bottom tracking, (2) GPS GGA and
105 (3) GPS VTG. Note that GGA and VTG refer to NMEA-0183 protocols for outputting GPS
106 instrument position, quality and velocity information(1). GPS data are acquired at up to 10 Hz
107 and read directly by the ADCP using the NMEA-0183 standard protocol. All three track
108 references are logged for each transect, with the most accurate one selected during post-
109 processing. The ADCP uses the East, North, Up (ENU) coordinate system because it allows for
110 free movement and rotation of the ADCP with respect to the cableway orientation. Orientation
111 of the ADCP is measured using a magnetic compass which was regularly calibrated on site.

112 The ADCP data were collected and post-processed using SonTek® RiverSurveyor Live software.
113 The post-processing consisted of verifying proper settings, including applying a site specific -
114 29.4 magnetic declination compensation to align to magnetic north, a constant 0.06 m
115 transducer depth offset, choosing the appropriate track reference, depth reference, and
116 screening for poor quality data. When GPS Quality = 4, indicating RTK positioning, GPS GGA
117 was selected as the track reference. GPS GGA was selected over bottom tracking due to
118 occasional anomalous samples and the unknown effects of bottom tracking on ice. For GPS
119 quality below Quality = 4, the GPS parameters were evaluated further and the track reference
120 was chosen per GPS quality, HDOP, and the number of logged satellites. Next, width and
121 velocity were evaluated with the selected track reference. Transects were removed if width was
122 an outlier or if a significant number of velocity profiles were missing. Velocity vectors were
123 analyzed visually for uniform and homogenous flow. If velocity vectors indicated significant
124 non-uniform flow, the measurement was excluded (**Table S1**).

125 The ADCP assumes homogenous flow therefore all beams must measure the same velocity
126 field. If one or more beams separate from the others, velocity and tracking ability of the system
127 is compromised. Under uniform flow conditions the signal-to-noise-ratios (SNR) of all 4 beams
128 converge and trend together, showing the same signal decay throughout the water column.
129 Divergence of the SNR from one or more of the beams is an indication of beam separation. SNR
130 vs. depth plots were visually reviewed for each measurement to inspect for beam separation.
131 Occasional beam separation is to be expected in turbulent flow environments and was
132 considered acceptable; however, transects with consistent, significant beam separation were
133 excluded.

134 The ADCP collects depth using an independent vertically-oriented sensor, as well as an average
135 depth recorded by four angled velocity beam sensors used for bottom tracking. The vertical
136 beam depth was used as the primary depth reference in this study because: (1) it is oriented

137   flush with M9 face; (2) The four velocity beams are oriented slanted relative to the face of the
138   ADCP and thus also the channel bed; (3) Depth from the velocity beams is calculated as the
139   average depth from the 4 beams, which can sometimes overly smooth bathymetry; (4) The
140   cableway-tethered Hydroboard provides a stable platform which helps minimize the effect of
141   vertical beam sensor tilt, which is not otherwise compensated for by RiverSurveyor Live
142   software; (5) The ADCP's vertical beam is wider than the velocity beams, signifying there would
143   be no significant change in measured depth given that sensor tilt angles were small. For these
144   five reasons, the vertical beam depth was chosen as the primary depth reference. When data
145   are missing from one depth reference the ADCP fills in missing data using the previous depth
146   reference. If the vertical beam reports an anomalous value while in transect, the bottom track
147   was chosen for the depth reference. For each profile, cross sectional area was evaluated and
148   transects with outliers were excluded.

149   The ADCP has a minimum depth range of approximately 30 cm, therefore the shallow margins
150   of the channel cannot be directly measured. The RiverSurveyor Live software uses stationary
151   edge measurements collected on each bank to estimate the discharge of the unmeasured
152   portion of the channel. The stationary edge measurements included manually estimating the
153   distance from the edge of the water to the ADCP on site, and collecting a minimum of 10
154   velocity profiles at the minimum readable depth. While post processing, transects with no edge
155   data on both banks were removed. Transects with reliable edge data on only one bank were
156   removed when other transects within a measurement hour contain accurate edge data. In this
157   reach of the Rio Behar, most of the flow was carried through the central portion of the channel
158   where reliable depth and velocity data were collected – thus we deem uncertainty due to
159   limited velocity cells at the edge of each profile to be minimal.

160   Each measurement time was manually copied from a Matlab file generated by River Surveyor
161   Live. The time is stored in the GPS.Utc structure array formatted as hhmmss.s. No data
162   GPS.Utc values were ignored and the minimum recorded time from the first profile collected in
163   each hour was assigned as the measurement start time, and the maximum recorded time from
164   the last profile collected in each measurement hour was assigned as the measurement end
165   time. After manual quality control checks and removal of anomalous transects, the mean
166   discharge and standard deviation were calculated for remaining transects in each measurement
167   hour to yield a final, averaged discharge measurement and associated measurement standard
168   deviation (**Table S1**).

169

170

171 **Supplementary Data Table S1**: Quality-controlled Acoustic Doppler Current Profiler (ADCP)
172 measurements of supraglacial river discharge and associated standard deviations collected in
173 the Rio Behar (67.049598N, -49.0201453W).

Hour	Date	Start Time (utc)	End Time (utc)	Mean Discharge (Q) $m^3 s^{-1}$	Standard Deviation Q $m^3 s^{-1}$	Number ADCP Profiles Used	Number ADCP Profiles Collected
1	7/20/2015	1:16:34 PM	1:24:59 PM	8.17	0.33	2	4
2	7/20/2015	1:58:58 PM	2:22:51 PM	7.14	0.61	2	6
3	7/20/2015	3:09:28 PM	3:22:43 PM	7.98	0.08	4	4
4	7/20/2015	4:10:00 PM	4:21:30 PM	9.96	0.22	4	4
5	7/20/2015	5:29:44 PM	5:39:04 PM	13.77	0.21	4	4
6	7/20/2015	6:24:16 PM	6:27:28 PM	17.31	0.95	3	4
7	7/20/2015	7:34:21 PM	7:45:52 PM	21.33	*	1	4
8	7/20/2015	8:41:43 PM	8:55:11 PM	24.87	0.89	3	4
9	7/20/2015	9:43:09 PM	9:54:51 PM	26.73	0.76	4	4
10	7/20/2015	10:41:51 PM	11:03:06 PM	25.50	0.50	3	6
11	7/20/2015	11:42:26 PM	11:54:32 PM	22.80	0.34	4	4
12	7/20/2015	12:45:08 AM	12:54:42 AM	19.60	0.45	5	5
13	7/20/2015	1:48:29 AM	1:56:33 AM	15.66	0.86	2	3
14	7/21/2015	2:41:31 AM	2:49:30 AM	11.92	0.34	4	4
15	7/21/2015	3:06:19 AM	3:13:42 AM	10.15	0.51	3	4
16	7/21/2015	4:00:02 AM	4:06:54 AM	8.63	0.71	3	4
17	7/21/2015	5:06:27 AM	5:14:59 AM	6.01	0.36	3	4
18	7/21/2015	5:53:43 AM	6:05:30 AM	6.25	0.35	2	4
19	7/21/2015	6:55:17 AM	7:13:10 AM	5.38	0.23	4	4
20	7/21/2015	7:48:33 AM	8:10:16 AM	4.72	0.69	3	6
21	7/21/2015	9:15:21 AM	9:26:41 AM	4.61	0.23	3	4
22	7/21/2015	10:11:00 AM	10:18:30 AM	5.10	0.18	3	4
23	7/21/2015	11:10:00 AM	11:16:30 AM	5.46	0.19	4	4
24	7/21/2015	12:11:40 PM	12:21:00 PM	6.00	0.10	4	4
25	7/21/2015	12:59:20 PM	1:08:20 PM	6.59	0.08	4	4
26	7/21/2015	2:06:50 PM	2:16:20 PM	7.05	0.06	3	4
27	7/21/2015	3:10:30 PM	3:19:30 PM	7.89	0.13	4	4
28	7/21/2015	4:09:20 PM	4:21:30 PM	9.00	0.54	5	6
29	7/21/2015	5:10:50 PM	5:22:30 PM	12.65	0.19	4	4
30	7/21/2015	6:05:40 PM	6:15:40 PM	16.15	0.091	3	4
31	7/21/2015	6:57:10 PM	7:07:20 PM	19.28	0.70	3	4
32	7/21/2015	8:04:20 PM	8:14:20 PM	22.16	0.693	4	4
33	7/21/2015	8:59:30 PM	9:07:10 PM	23.9	0.404	2	4
34	7/21/2015	10:00:20 PM	10:08:30 PM	25.54	0.771	4	4
35	7/21/2015	10:55:10 PM	11:03:30 PM	25.2	*	1	4
36	7/22/2015	12:01:58 AM	12:11:09 AM	21.7	0.857	3	4
37	7/22/2015	12:58:15 AM	1:05:30 AM	19.06	0.452	4	4

38	7/22/2015	2:03:15 AM	2:12:15 AM	16.38	0.398	4	4
39	7/22/2015	3:15:19 AM	3:23:24 AM	14.1	0.173	4	4
40	7/22/2015	4:06:03 AM	4:12:43 PM	11.82	0.252	4	4
41	7/22/2015	5:01:09 AM	5:06:20 AM	10.78	0.292	3	4
42	7/22/2015	6:09:16 AM	6:19:04 AM	9.208	0.269	4	4
43	7/22/2015	7:08:18 AM	7:16:08 AM	8.585	0.488	4	4
44	7/22/2015	8:32:49 AM	8:39:51 AM	7.856	0.108	4	4
45	7/22/2015	8:53:01 AM	8:59:09 AM	7.578	0.092	4	4
46	7/22/2015	10:42:50 AM	10:50:58 AM	7.434	0.119	4	4
47	7/22/2015	11:09:01 AM	11:16:03 AM	7.566	0.155	4	4
48	7/22/2015	12:06:39 PM	12:13:32 PM	7.698	0.087	4	4
49	7/22/2015	1:42:29 PM	1:52:14 PM	8.647	0.211	4	5
50	7/22/2015	2:10:39 PM	2:18:07 PM	9.352	0.349	4	4
51	7/22/2015	3:09:29 PM	3:18:30 PM	12.11	0.144	4	4
52	7/22/2015	4:00:17 PM	4:10:12 PM	14.3	0.384	4	4
53	7/22/2015	5:06:23 PM	5:13:14 PM	17.99	0.798	4	4
54	7/22/2015	6:05:37 PM	6:13:15 PM	20.63	0.849	4	4
55	7/22/2015	7:00:30 PM	7:07:41 PM	22.06	0.539	4	4
56	7/22/2015	8:00:10 PM	8:06:52 PM	21.99	0.275	4	4
57	7/22/2015	9:05:02 PM	9:12:32 PM	21.75	0.442	4	4
58	7/22/2015	9:59:52 PM	10:09:40 PM	20.58	0.393	4	4
59	7/22/2015	11:01:57 PM	11:09:26 PM	19.76	0.209	4	4
60	7/23/2015	12:03:02 AM	12:09:31 AM	17.68	0.401	4	4
61	7/23/2015	1:04:40 AM	1:10:59 AM	15.88	0.147	4	4
62	7/23/2015	1:59:57 AM	2:05:26 AM	13.36	0.19	4	4
63	7/23/2015	3:02:19 AM	3:08:15 AM	10.01	0.147	4	4
64	7/23/2015	3:59:34 AM	4:05:21 AM	9.525	0.386	4	4
65	7/23/2015	5:00:03 AM	5:13:52 AM	8.672	0.222	2	6
66	7/23/2015	6:19:54 AM	6:28:09 AM	6.65	0.147	3	4
67	7/23/2015	7:00:43 AM	7:08:35 AM	7.198	0.007	3	4
68	7/23/2015	8:15:43 AM	8:23:19 AM	6.545	0.264	2	4
69	7/23/2015	9:00:05 AM	9:09:22 AM	7.044	0.159	4	4
70	7/23/2015	10:08:58 AM	10:16:04 AM	6.104	0.053	3	4
71	7/23/2015	11:00:27 AM	11:07:41 AM	6.128	0.08	4	4
72	7/23/2015	12:05:03 PM	12:12:11 PM	6.586	0.082	4	4

* standard deviation not calculated because only one ADCP profile remains after manual quality control and data filtering

174

175

176

177

1.3 Field measurements of ice surface lowering (ablation):

179 Ice surface ablation measurements:

180 Repeated ice surface lowering measurements were collected at fifteen bamboo ablation stakes
181 distributed around our ADCP ice camp (black star, **Figure 1**). Fifteen stakes were placed at
182 random distance/direction pairs from a common center to capture spatial variation in ablation.
183 Thirteen stakes were in weathering crust which was the typical ice surface at the site. One
184 stake was placed in a remnant patch of seasonal snow underlain by solid ice, and one in a patch
185 of solid bare ice covered in dispersed cryoconite debris. Stakes were drilled at least 1 m deep
186 into the surface and left to freeze in for 24 hours before initiating measurements. Prior to each
187 measurement, a 24 x 24 cm wooden "ablation board" was placed at the foot of each stake
188 (**Figure S2a**). The board was oriented to true north and measurements were made from a
189 marked point on the base of the board to a marked point at the top of the stake. This protocol
190 minimized errors due to local variations in ice sheet surface micro-topography or shifts in
191 measurement datum. Careful inspection of each stake was made prior to each measurement.
192 To ensure accurate surface lowering measurements, all stakes were pressed firmly downward
193 prior to measurement to ensure that any loosening stakes were properly seated in the bottom
194 of their drill holes.   Changes in ice sheet surface height were measured at sub-daily intervals
195 (~3 to 12 hours). A total of thirteen, sub-daily surveys of the heights of each stake above the
196 ice surface were conducted between 1100 July 21 and 1415 July 24, 2015. Data from all 15
197 stakes were averaged together and their standard deviations computed to mitigate for known
198 spatial variability in ablation stake courses (2).

199 Approximately 9.1 km east of our base camp, hourly changes in ice surface height were
200 obtained from a downward-looking Campbell Scientific SR50A sonic ranger fitted on the
201 KAN_M automatic weather station (AWS), located just outside the 2015 Rio Behar catchment
202 boundary (67.0667 N, 48.8327 W, 1270 m a.s.l., see **Figure 1**). These data have at least ~1 cm
203 vertical uncertainty owing to pressure-transducer correction and sensor noise. Use of raw
204 surface lowering data from this sensor and our ablation stakes is appropriate for estimation of
205 *M* during our field experiment in late July, as there was no new snow accumulation and
206 negligible refreezing or sublimation in above-zero air temperatures. Meteorological data from
207 the KAN_M AWS were also used to supply the necessary inputs to run the point SEB model (see
208 **SI**). The KAN_M AWS is maintained by the Geological Survey of Denmark and Greenland
209 (GEUS), as part of the Greenland Analogue Project and the Programme for Monitoring of the
210 Greenland Ice Sheet (www.PROMICE.dk).

211 To compare climate/SMB model outputs of melt *M* (mm liquid water) directly with the
212 described surface lowering data, model outputs of *M* were converted to units of solid ice
213 equivalent (mm). To provide associated uncertainty ranges on these conversions, two different

214   ice densities were assumed: 0.688 g cm^{-3} (the average near-surface ice density from ten

215   shallow cores obtained at the Rio Behar base camp site, from Cooper, *et al.* (3)) and 0.918 g cm^{-}

216   ^{3} (density of pure solid ice) to provide an upper and lower bound of ice surface lowering,

217   respectively. The resultant ranges of modeled ice surface lowering due to *M* are presented

218   alongside field measurements in **Figure 7b**.
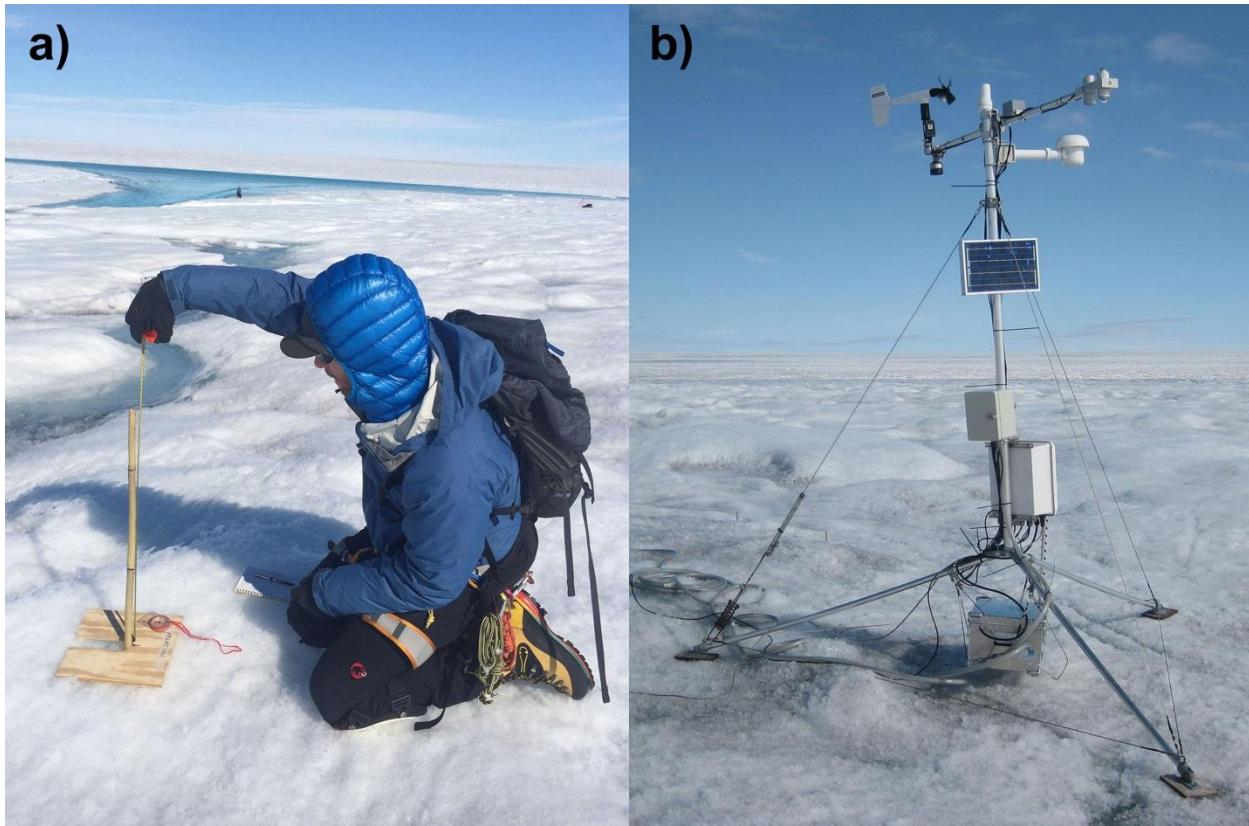
219

220   **Figure S2: Surface lowering data were collected from (a) manual steel tape measurements at fifteen**

221   **bamboo ablation stakes distributed around Rio Behar base camp; and (b) sonic ranging data from the**

222   **PROMICE KAN_M automated weathering station maintained by the Geological Survey of Denmark**

223   **and Greenland (GEUS). Base camp and KAN_M station locations are shown in Figure 1. Photos by (a)**

224   **Åsa K. Rennermalm and (b) Dirk van As.**

225

226   **2. Remote sensing data**

227   Remotely sensed datasets consisted mainly of high resolution WorldView-1/2 satellite imagery

228   and associated digital elevation models (DEMs) derived from stereo-photogrammetry, and RGB

229   camera images acquired from a custom built fixed-wing UAV(4) (drone). Archived WorldView

230   satellite data were also used to test our SUH model against two previous field studies (5, 6). A

231   previously published map (7) of supraglacial internally drained catchments (IDCs) derived from

232   a 19 August 2013 panchromatic Landsat-8 image supplied the 799 IDC boundaries used in this
233   study, and daily MODIS albedo retrievals (MYD10A1 product) used to validate (RACMO2.3) or
234   drive (HIRHAM5) climate/SMB albedo. A summary table of these data, including product ID
235   numbers, acquisition dates, and what they were used for is presented in **Table S2.**

236   **Table S2: Remotely sensed datasets used in this study**

Dataset type	ID	Spatial resolution	Acquisition dates	Purpose
WorldView-1 satellite imagery	102001004202CD00	0.5 m	18 July 2015	Detailed mapping of supraglacial hydrologic features (rivers, lakes, moulins, channel heads), and stereo-photogrammetric DEM generation
	1020010043165100	0.5 m	18 July 2015	Stereo-photogrammetric DEM generation and catchment boundary extraction
	10200100354A5700	0.5 m	29 October 2014	
	102001003434B00	0.5 m	29 October 2014	
	102001003376CE00	0.5 m	19 September 2014	
	10200100318C0D00	0.5 m	19 September 2014	
	1020010008AB4800	0.5 m	15 July 2009	McGrath *et al.*(6) SUH case study
	103001000CB46800	0.5 m	12 July 2011	Chandler *et al.* (5) SUH case study
WorldView-2 satellite imagery	1030010046354000	2.0 m	17 July 2015	Mapping snow covered area; Calculating supraglacial lake depth; Comparing small streams in concurrent WV and UAV images; Supporting and validating supraglacial hydrologic feature mapping
	1030010045092500	2.0 m	18 July 2015	
	1030010047C39F00	2.0 m	18 July 2015	
	1030010046A5F200	0.5 m & 2.0 m	24 July 2015	
	10300100470EB600	0.5 m & 2.0 m	24 July 2015	
WorldView-3 satellite imagery	104001000EB35500	0.5 m & 2.0 m	21 July 2015	
UAV imagery	-	0.25 m	20-22 July 2015	
MODIS albedo	MYD10A1 daily product	500 m	20-21 July 2015	Validate or drive climate/SMB models
SETSM DEM	ArcticDEM_15_39_5_2	2.0 m	30 August 2011	Assist catchment boundary delineation
SPIRIT DEM	DRONNING_INGRID_080627	40.0 m	27 June 2008	

237

238

239 2.1 Satellite image and digital elevation model (DEM) products:

240 A total of eight 0.5 m panchromatic images from the WorldView-1 and WorldView-2 satellites
241 were used to construct stereo-photogrammetric DEMs and map supraglacial streams, rivers,
242 moulins, and channel heads (incision initiation) within the Rio Behar catchment. Three 2.0 m
243 multispectral images from WorldView-2 were used to map snow areas. Archived WorldView-1
244 images were used to map the IDCs of two previously published field studies (5, 6). All
245 WorldView satellite images were acquired through the Polar Geospatial Center (PGC) and were
246 orthorectified using the satellite positioning model (also known as the rational function model)
247 and projected into a polar stereographic coordinate system using PGC code
248 (https://github.com/PolarGeospatialCenter/imagery_utils). One scene of Stereo-
249 Photogrammetric Digital Elevation/Surface Models (SETSM, spatial resolution 2 m) released by
250 the PGC and the Byrd Polar Research Center Glacier Dynamics Group
251 (http://www.pgc.umn.edu/elevation/stereo)and one scene of SPOT 5 stereoscopic survey of
252 Polar Ice: Reference Images and Topographies (SPIRIT, spatial resolution 40 m) DEM released by
253 the International Polar Year (IPY) project
254 (https://theia.cnes.fr/rocket/#/search?collection=Spirit) were also used to aid extraction and
255 evaluation of the Rio Behar catchment boundary.

256

257 2.2. Unmanned Aerial Vehicle (UAV):

258 To supplement and verify the WorldView satellite data acquired during the field experiment, a
259 fixed-wing UAV as described by Ryan *et al.* (4) acquired aerial RGB camera imagery over Rio
260 Behar catchment from 20-22 July as part of three successive surveys beginning at 12:04 UTC on
261 20 July 2015 and finishing at 18:45 UTC on 22 July 2015. This UAV has a wingspan of 2.12 m and
262 is powered by eight custom-made (14.4V) lithium-ion battery packs. Propulsion is provided by a
263 715W brushless electric motor which turns a 12 x 8 inch foldable propeller. Autonomous
264 control is provided by a Pixhawk autopilot designed by the PX4 open-hardware project and
265 manufactured by 3D Robotics (https://pixhawk.org/modules/pixhawk). The Pixhawk utilizes a
266 single-frequency GPS, gyroscope, accelerometer, magnetometer and barometer for flight
267 control. With this configuration, the UAV has a cruising speed of 60 km/hour and endurance of
268 1.5 hours, allowing the UAV to comfortably fly survey missions of up to 80 km. A total of 3,795
269 overlapping images at an altitude of 1,000 m above the ice surface were obtained over a 102
270 km² area extending beyond the Rio Behar catchment boundaries. These camera images were
271 processed with Agisoft PhotoScan Pro (8) and an ortho-mosaic with a 30 cm ground sampling
272 distance (GSD). A digital elevation model with a 1 m GSD was also generated but not used in
273 the study. Georeferencing accuracy of orthorectified products was aided by installation of four
274 GPS surveyed (single-frequency, accurate to 3 m horizontally and up to 15 m vertically) red

275 tarpaulins at four distributed locations around Rio Behar catchment (position 1: 67.0486°N, -
276 49.0208°E, 1207.5 m; position 2: 67.0610°N, -48.9025°E, 1289.5 m; position 3: 67.0315°N, -
277 48.9582°E, 1248.1 m; position 4: 67.0214°N, -48.9281°E, 1271.7 m) for use as ground control
278 points (GCP, red plus symbols, **Figure 1**). The primary utility of these higher-resolution data was
279 to supplement WorldView satellite data for the purpose of visually identifying small areas of
280 confirmed and potential catchment leakage (i.e. internal subcatchments) draining to small
281 internal moulins, and crevasse fields. Comparison of 0.25 m UAV imagery with panchromatic
282 WorldView-1 imagery confirms good detection of supraglacial streams, even down to sub-meter
283 widths (**Figure S3**).

284
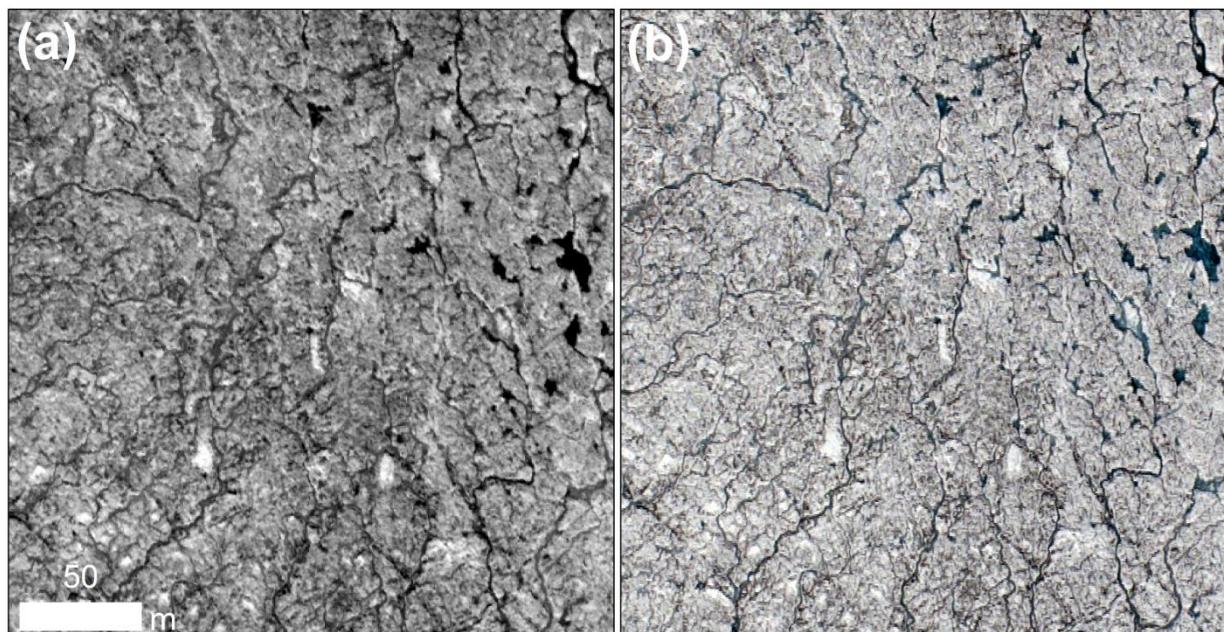
285 **Figure S3. WV1 and UAV discrimination of small supraglacial streams. Concurrent (a) 0.5 m**
286 **panchromatic WorldView-3 (WV3) image (catalog ID: 104001000EB35500) and (b) 0.25 m UAV image**
287 **acquired on 21 July 2015 for a headwater area of the Rio Behar catchment, showing that the same**
288 **small streams can be discerned in both WV3 and UAV imagery. Image center locations are 67.079 N,**
289 **48.916 W and image sizes are 250 × 250 m. WV3 imagery Copyright 2015 DigitalGlobe, Inc.**

290

291

292 **3. Remote sensing data processing**

293 High resolution satellite and UAV mappings of Rio Behar catchment bookend and overlap with
294 the ADCP river discharge and ice ablation measurements. These remotely sensed geospatial
295 data were post-processed to produce precise maps of Rio Behar catchment boundaries,
296 supraglacial drainage patterns, small internal moulins (leakage), and snow cover as follows:

297 <u>3.1 processing of stereo-photogrammetric DEMs from WorldView satellite imagery:</u>

298 High resolution stereo-photogrammetric DEMs were derived from along-track stereo
299 WorldView panchromatic satellite imagery using the open source Ames Stereo Pipeline (ASP)
300 toolkit methods (9, 10). Images were resampled to 1 m resolution before processing to reduce
301 computational time needed for DEM production. The output of ASP consists of point clouds
302 that are spatially filtered to produce 3 m posting DEMs. Horizontal positioning accuracy is
303 typically better than 5 m. In this study, the derived DEMs were used to extract Rio Behar
304 catchment boundaries (**Figure 1**) following hydrologic analysis as follows: 1) only the
305 topographic depression at the catchment outlet was used as meltwater sink and a partially
306 filled DEM raster was created (11); 2) flow direction was identified by using this partially filled
307 DEM; 3) the catchment boundary was extracted using the Basin function in ArcGIS software. A
308 total of five digital elevation models (DEMs) constructed for Rio Behar catchment over the
309 period 2008-2015 confirm overall long-term stability of the topographic boundaries of this
310 particular IDC with interannual area variations of just 1.9 – 3.3 % (**Figure 1**).

311 <u>3.2 Image processing of WorldView satellite imagery</u>:

312 Supraglacial stream/river networks were delineated from the 0.5 m panchromatic 18 July 2015
313 WorldView-1 image, following the method of *Yang et al.* (12). Variable ice surface backgrounds
314 were first eliminated by spectral analysis and non-local means denoising; small supraglacial
315 rivers were enhanced by Gabor filtering; and continuous supraglacial river networks were
316 obtained by path opening (12). Next, a global threshold of 120 was used to classify the original
317 WorldView-1 image (gray value ranges from 0 to 503) to extract supraglacial lakes. Finally, a
318 meltwater mask raster was generated by combining the described river and lake binary masks.

319 In total, some 3380.6 km of supraglacial stream/river lengths were mapped in the optimal Rio
320 Behar catchment boundary, yielding a drainage density of 53.6 km/km^2. The mean channel
321 width was 2.4 ± 1.5 m and surface meltwater covered 8.2 % of the ice surface. Two WorldView-
322 1 panchromatic images (18 July 2015 and 24 July 2015) and repeated UAV sorties (20-22 July
323 2015) confirmed that the supraglacial river network and four small, interconnected supraglacial
324 lakes (**Figure 1**) remained intact and actively flowing throughout the 20-23 July 2015 field
325 experiment.

326 Uppermost headwater channel heads (initiation points) of first-order tributaries were mapped
327 visually from the 0.5 m panchromatic 18 July 2015 WorldView-1 image (catalog ID:
328 102001004202CD00, see **Table S2**), for both Rio Behar catchment headwaters (total 839
329 mapped, termed "inner" channel heads, **Figure 1**) and adjacent catchments (total 780 mapped,
330 termed "outer" channel heads, **Figure 1**). Connecting these inner and outer channel heads
331 yields minimum and maximum plausible extents of the Rio Behar catchment, respectively. In

332   addition, 49 confirmed and 24 possible internal moulins and/or moulin complexes were
333   manually identified from the 18 July WorldView-1 image cross-checked by the 20-22 July 2015
334   UAV camera images and the 24 July 2015 WorldView-1 panchromatic image.

335   Connecting the 780 outer channel heads for the adjacent catchment provides the maximum
336   plausible catchment extent (area 69.1 km^2, all pixels lying outside this boundary confidently do
337   not flow to the Rio Behar terminal moulin), whereas connecting the 839 channel heads for the
338   Rio Behar catchment provides a conservative minimum plausible extent (51.4 km^2, with all
339   pixels inside this boundary and lying outside of the small internally drained subareas
340   confidently flowing to the terminal moulin) for the Rio Behar IDC during the field experiment.
341   An intermediate, "optimal" catchment boundary was derived using the 18 July 2015
342   WorldView-1 stereo-photogrammetric DEM, adjusted for small areas of stream "breaching"
343   (piracy) across topographic ice divides (12, 13) and small areas subareas of internal moulin
344   drainage (**Figure 1**). Visual inspection of WorldView and UAV imagery along this topographic
345   boundary, revealed small areas undoubtedly draining away from Rio Behar catchment (but
346   included inside its topographic divide) owing to stream channel breaching of the topographic
347   divide. These small areas (totaling 2.7 km^2) were manually eliminated from the 18 July 2015
348   DEM boundary. Similarly, small areas undoubtedly draining into the Rio Behar catchment (but
349   outside the topographic divide) were manually added (totaling 0.8 km^2). Finally, small
350   subcatchments flowing to internal moulins (total area 1.6 km^2) and crevasse fields (4.1 km^2)
351   were eliminated. Crevassed areas were determined visually in WorldView and UAV imagery
352   and eliminated from our lower-bound estimate of watershed area, so their potential storage
353   falls within the central and lower bound of modeled runoff uncertainty in **Figure 3** and **Figure
354   7a**. We submit that the resultant optimal catchment delineation (area 63.1 km^2) combines the
355   strengths of high-resolution remote sensing of the supraglacial stream/river drainage pattern
356   with topographic divides from a simultaneous WorldView stereo-photogrammetric DEM.

357   Small subcatchments draining to the "confirmed" and "possible" internal moulins were
358   identified in the WorldView and UAV imagery and also removed to demarcate the conservative
359   Rio Behar catchment area estimate (**Figure 1**). To do this, subcatchments draining into
360   confirmed and potential internal moulins were labeled with unique IDs in ArcGIS. These labeled
361   river networks were then used as seed regions for region expansion, by applying the path
362   distance allocation function in ArcGIS to calculate the nearest source for each seed river
363   network based on the minimum cumulative cost over the cost surface as per Yang and Smith
364   (7). The resultant allocation map partitioned Rio Behar catchment into three different
365   categories, with two categories, subcatchments draining into confirmed (total area 1.6 km^2)
366   and potential (area 0.7 km^2) internal moulins eliminated to yield the conservative catchment
367   boundary (area 51.4 km^2).

368   3.3 Estimation of supraglacial lake volumes:

369   Four supraglacial lakes are integrated into the stream/river network of Rio Behar catchment

370   (**Figure 1**). To investigate whether water impoundment in these lakes could explain the

371   observed deficit between ADCP and climate/SMB model estimates of R, we used multi-spectral

372   WorldView-2 images acquired on 18 July 2015 (catalog ID: 1030010045092500 and

373   1030010047C39F00) and 24 July 2015 (catalog ID: 1030010046A5F200) to calculate their

374   respective volume changes over the time of our ADCP field experiment. We used the method

375   of Pope *et al.* (14), which entails building a DEM of the supraglacial lake basins when they are

376   empty, then intersecting the DEM with a remotely sensed lake mask when the basin is occupied

377   with water. DEM elevations within the shorelines of the lake mask are then summed to obtain

378   lake volume. In the 18 July 2015 WV DEM, all four depressions are partially occupied by lakes

379   and therefore could not be used as the base DEM. We built a new DEM from the preceding fall,

380   using a stereo WorldView-1 image pair acquired on 29 October 2014 (catalog ID:

381   1020010034334B00 and 10200100354A5700). In this post-melt stereo image pair, all four lake

382   depressions are empty and were used to estimate supraglacial lake volumes the following year

383   (topographic depressions are largely controlled by bedrock(11) and are therefore deemed

384   stable over the period October 2014 to July 2015). We extracted lake masks using a band ratio

385   of Band 2 (blue, 450-510 nm) to Band 8 (near infrared, 860-1040 nm), with a global ratio

386   threshold set to 1.25 following the method of Smith *et al.* (13).   The two resultant lake masks

387   were used to clip DEM-modeled topographic depressions and the topographic depression

388   volumes located in the lake masks were calculated as the lake volumes.

389   During 18-24 July2015, all four supraglacial lakes shrank, for a total volume reduction of 2.2 ±

390   4.3×10^{-4} km^3. Divided over by our optimal catchment area, this lake volume change equals to

391   3.5 ± 6.8 mm runoff in the entire Rio Behar catchment, with the uncertainty estimate computed

392   as one standard deviation from the mean in lake elevation along the shoreline as per Pope *et al.*

393   (14). Interpolating linearly between the two satellite acquisition dates, this corresponds to a

394   water release of 0.02 ± 0.05 mm/hour, which is 3.2 ± 6.3 % of the average ADCP hourly

395   discharge (0.75 mm/hour). We therefore conclude that water impoundment in these four

396   supraglacial lakes cannot explain the observed runoff deficit between ADCP measurements and

397   climate/SMB models. Instead, they released a minor amount (<10 %) of meltwater to the

398   observed ADCP river discharge, lending further conservatism to our finding of climate/SMB

399   model overestimation of R.

400   3.4 Snow classification:

401   The images acquired during our field experiment were classified to determine the fractional

402   area of snow cover during the time of our ADCP measurements. These images consisted of our

403   UAV camera image mosaic (20-22 July 2015, RGB bands, spatial resolution 0.3 m) and three

404
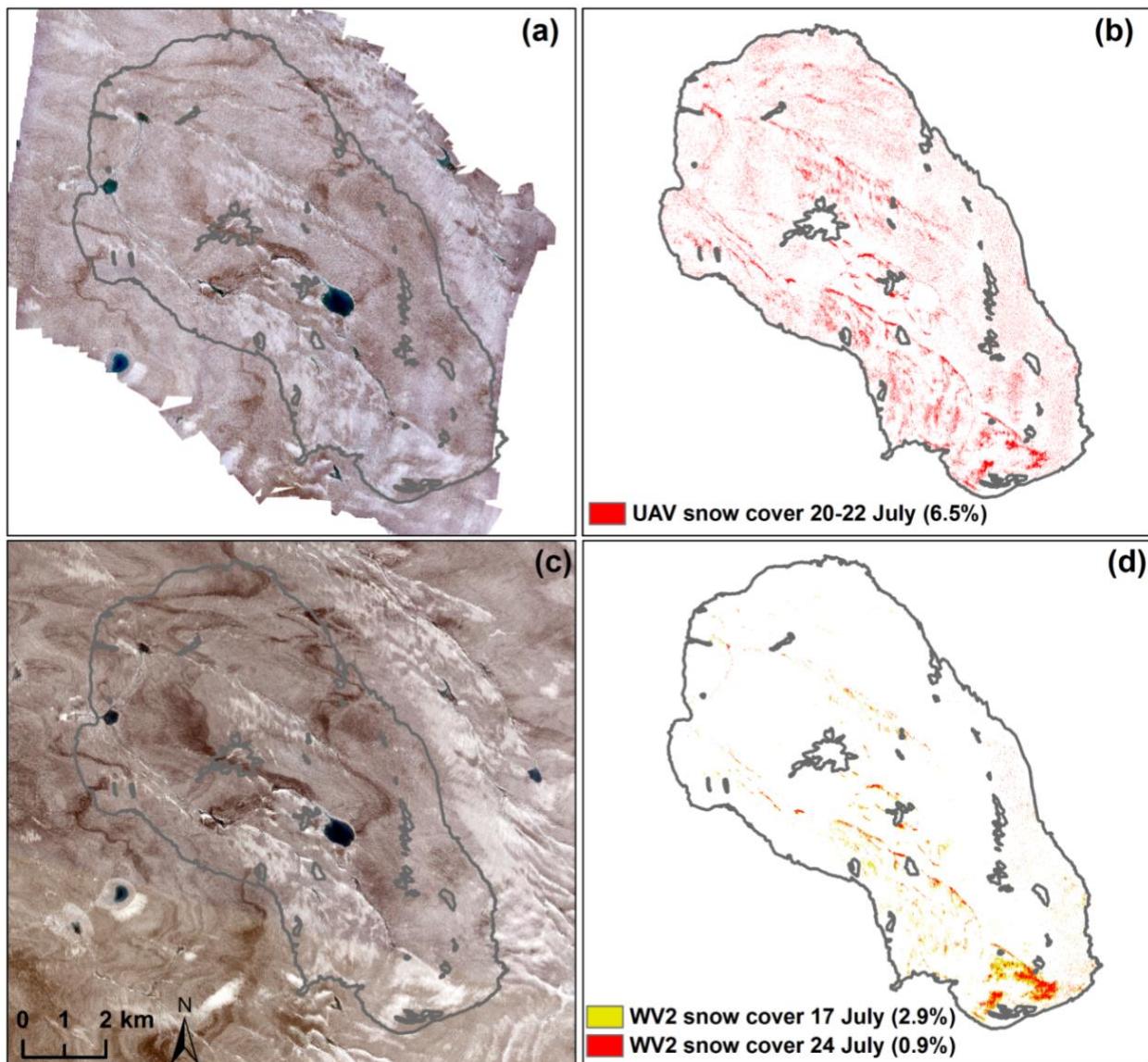
Figure S4: Snow cover classifications derived from a UAV image mosaic acquired 20-22 July 2015,
(RGB bands, spatial resolution 0.3 m) and two WorldView-2 multi-spectral images acquired 17 and 24
July 2015 (catalog IDs: 1030010046354000 and 1030010046A5F200, 8 bands, 2.0 m): (a) UAV image
mosaic, (b) UAV snow classification, (c) 24 July WorldView-2 multi-spectral image, (d) snow
classifications of 17 July and 24 July WorldView-2 images. Viewed collectively, these maps confirm
that Rio Behar catchment was largely snow-free during the field experiment, confirming climate/SMB
model assumptions of bare-ice conditions and ruling out temporary meltwater retention in seasonal
snow/firn as an explanation for the observed overestimation of modeled runoff *R*. WorldView-2
imagery Copyright 2015 DigitalGlobe, Inc.

WorldView-2 multi-spectral images (one from 17 July 2015 and two from 24 July 2015, 8 bands,
2.0 m resolution). First, a supervised k-Nearest Neighbors (k-NN) algorithm from the scikit-
learn Python module (15) was used to classify the UAV image. Next, a supervised maximum

17

417 likelihood algorithm from the ArcGIS software was used to classify both the UAV and WV2

418 images. The two classification algorithms were trained by manually digitizing areas of snow,

419 bare ice and surface water. Once trained, pixels were classified into one of the three classes

420 (snow, ice or water) and three snow classification maps were obtained (**Figure S4**). The

421 resultant snow cover fractions are 6.5% in the 20-22 July UAV mosaic, and 2.9% and 0.9% in the

422 17 and 24 July WV2 images, respectively. This difference between the two sensors is attributed

423 to the higher spatial resolution of the UAV data (0.25 m, versus 2.0 m for multi-spectral WV2)

424 allowing better mapping of small snow patches, the use of near-infrared channels on WV2 (UAV

425 is visible-spectrum only), or both. For the purpose of this study, both UAV and WorldView-2

426 snow classifications indicate that snow cover was largely absent from the Rio Behar catchment

427 during the time of our field experiment. This confirms the climate/SMB model assumptions of

428 bare-ice conditions at the time of our field experiment, and largely rules out meltwater

429 retention in seasonal snow as a significant physical mechanism for the observed discrepancy

430 between observed and climate/SMB modeled runoff *R*.

431

432

433 **4. Supraglacial hydrograph analysis**

434 The 72-hour ADCP time series of supraglacial river discharge Q comprises the core field dataset

435 of this study and was processed using traditional hydrograph analysis methods for terrestrial

436 watershed hydrology. These steps include quantification of hydrograph recession flow,

437 hydrograph separation, derivation of runoff coefficients/correction factors, derivation of the

438 Unit Hydrograph (UH) for Rio Behar catchment, and using the UH to calibrate a Synthetic Unit

439 Hydrograph (SUH) for broader extension across the southwest GrIS ablation zone, as follows:

440 4.1 Recession flow:

441 Of particular benefit to our hydrograph analysis was a nightly cessation of melt production from

442 approximately midnight (0000) to 0600 local time each day, and a lack of any precipitation

443 during the field experiment. The first yielded an unambiguous nightly period of hydrograph

444 recession (i.e. an interval of receding runoff from the catchment during which no new

445 climatological meltwater production occurred in the catchment). The second allowed all

446 variations in the discharge hydrograph *Q* to be attributed solely to the diurnal cycle in

447 climate/SMB modeled melt production *M*, without the added complication of rain-on-snow

448 events. Hydrograph recessions were observed nightly from approximately 0000 to 0600,

449 enabling quantification of exponential-decay hydrograph recession constants (*k*) and separation

450 of recession flow from falling limbs of the direct (observed) hydrograph. For the four models

451 supplying melt simulations *M* (i.e. all models except MERRA-2) we observed positive *M* values

452 (i.e. a "melt hyetograph") persisting for approximately 18 hours each day (from 0600 to 0000),
453 with cessation of melt production for approximately 6 hours each night (from 0000 to 0600).
454 During this nightly melt shutdown, supraglacial discharge (i.e. the direct hydrograph Q) receded
455 but did not terminate throughout the night. This signifies that 18 hours of melting on Rio Behar
456 catchment yields measurable runoff for at least 24 hours, and that hydrograph recession
457 constants must be obtained and used to separate direct vs. recession flow from each diurnal
458 cycle. We defined a 6 hour flow recession period from 0000 to 0600 nightly, for fitting with the
459 customary (16) exponential decay recession equation, $Q = Q_0 k^t$ (with Q_0 being the discharge at
460 recession initiation, Q the discharge at t hours later, and k the exponential-decay recession
461 constant). This fitting yielded k values of 0.79, 0.88, and 0.88 for July 21, 22, and 23,
462 respectively, then averaged to yield k=0.85 (averaging over multiple recessions is suitable for
463 characterizing mean recession behavior (17, 18)) for Rio Behar catchment over the period 21-23
464 July 2015. This is slightly lower than typical terrestrial catchment k values (>0.90) implying a
465 somewhat slower, "flatter" recession process but one not dissimilar from terrestrial catchments
466 (16, 19).

467 4.2 Hydrograph separation:

468 For each diurnal cycle, the described recession fitting equations were used to separate
469 recession flow from direct flow. A starting assumption is that for a single hour of melt
470 production, that hour may contribute to observed supraglacial river discharge for up to 24
471 hours. As such, residual meltwater produced during a typical 18 hour "melt-production day"
472 (i.e. from hour 0600 to 0000 local time) may arrive at the catchment's terminal moulin up to 42
473 hours later (i.e. from 0600 to 0000 of the following day). It is possible that recession flow
474 persists longer than 42 hours, but will likely be very small (<0.05 mm/h or <0.2% of the total
475 melt in one day) so a further time extension is not performed in this study. Because recession
476 flow extends into each following day, we derive two complete diurnal cycles of separated
477 hydrographs from the 72 hour record (**Figure S5**). The final outcome is thus two independent,
478 separated 42 hour hydrographs, obtained by subtracting recession discharge from the previous
479 day and retaining recession discharge into the following day. These two complete diurnal cycles
480 of separated hydrograph flow extend from 0600 July 21 to 0000 July 23 (termed July 21
481 hydrograph), and from 0600 July 22 to 0000 on July 24 (termed July 22 hydrograph), as derived
482 from hydrograph separation over the period 0000 July 21 to 0600 July 23 (**Figure S5**).

483
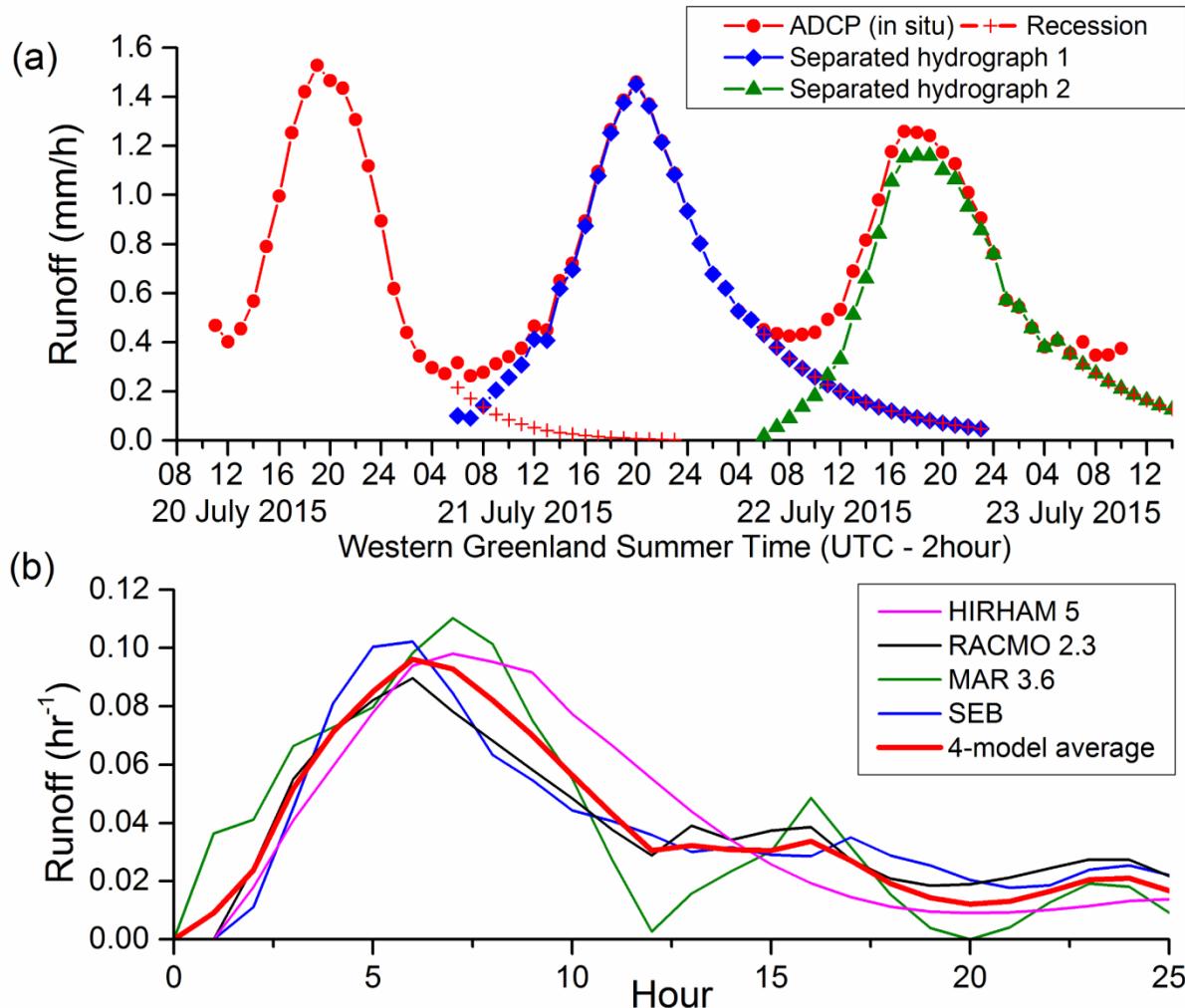
Figure S5: Rio Behar hydrograph separation and model-tuned synthetic unit hydrographs (UH) relating climate/SMB modeled melt to runoff delivered to the Rio Behar terminal moulin. (a) Hydrograph separation, yielding two separated diurnal hydrographs on 21/22 July and 22/23 July 2015. (b) Unit Hydrograph (UH) routing models calibrated to corrected climate/SMB melt outputs (M', see SI section 4.3) of the HIRHAM 5, RACMO 2.3, MAR 3.6, and SEB climate/SMB models. No UH is available for the MERRA-2 model because melt M is not an output of MERRA-2. The 4-model average (thick red line in Figure S5b) is used to generate theoretical 1 cm peak discharges in Figure 4b.

4.3 Runoff coefficients (correction factors) relating modeled M to observed R: Following derivation of these two separated hydrographs, runoff coefficients (c) may be calculated for each climate/SMB model as $c = Q/M$, where M is total "daily" (i.e. total 18 hour) catchment surface melt M as simulated by models, and Q the corresponding separated (i.e. total 42 hour)

498   associated catchment discharge. Note that in this particular study the derived coefficients
499   reflect inferred model over-prediction of R, so are more appropriately treated as empirical
500   correction factors for model output rather not quantification of physical runoff losses (e.g. to
501   infiltration, retention, missed crevasses etc.). The resultant c values ranges from 0.53-0.78,
502   indicating that 53-78% of the surface meltwater M simulated by models was measured as
503   physical runoff departing the ice surface to the Rio Behar catchment terminal moulin.   As a
504   practical step, these model-specific coefficients may be multiplied with model outputs of M to
505   yield a lower, empirically corrected estimate of runoff departing the ice surface. To
506   differentiate this derived quantity from the modeled variable runoff R we propose the term
507   "effective melt" (M') for this adjusted value of M, that is $M' = c * M$. These model-tuned runoff
508   coefficients / correction factors are supplied in **Table S3** and used as forcing data for the
509   instantaneous area-integrated runoff and SUH-routed computations of **Figure 5** and **Figure S11**.
510   Generation of runoff coefficients / correction factors and SUH parameters is not possible for
511   MERRA-2 because meltwater production M is not an output of this model. For ablation
512   surfaces possessing physical conditions like those sampled at Rio Behar catchment, these
513   coefficients may be multiplied by M to yield alternate, lower estimates of R in addition to
514   standard model output. Pending further study, they cannot be confidently extended to non-
515   similar surfaces or other times of the year. Future collection of supraglacial discharge
516   measurements across a range of ice surface type is needed to develop runoff coefficients /
517   correction factors and SUH parameters for other surfaces (in particular catchments containing
518   firn) and earlier/later times of year.

519   **Table S3: Rio Behar catchment runoff coefficients/correction factors and SUH parameters**

	Runoff coefficients / correction factors			Synthetic unit hydrograph (SUH) parameters*				
	July 21	July 22	Average	t_p (hr)	h_p (hr^{-1})	C_p	C_t	Gamma m
MAR 3.6	0.73	0.64	0.69	6.5	0.11	0.72	1.61	4.6
SEB	0.78	0.72	0.75	5.5	0.10	0.56	1.36	3.1
RACMO 2.3	0.62	0.67	0.65	5.5	0.09	0.49	1.36	2.1
HIRHAM 5	0.55	0.53	0.54	6.5	0.10	0.64	1.61	3.3
MERRA-2	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

520

521   * To apply SUH routing to a larger area of the ablating Greenland ice sheet surface, the above
522   values for C_p, C_t and m calibrated for Rio Behar catchment were applied to 799 surrounding IDCs
523   during the study period. Their catchment parameters t_p and h_p, however, were remotely sensed
524   for each individual IDC. For description of runoff coefficients / correction factors see Section
525   4.3. For descriptions of SUH parameters see Sections 5.1 and 5.2.

526 <u>4.4 Unit Hydrograph (UH) derivation:</u>

527 The unit hydrograph (UH) is a transfer function that is widely used for modeling catchment
528 runoff response to rainfall events for some unit duration and unit depth of effective water input
529 (i.e. the "excess" precipitation remaining and available to run-off following interception and
530 infiltration, analogous to our M') applied uniformly across the catchment(19). A "one hour UH",
531 for example, represents the characteristic response of a given catchment to a unit depth of
532 effective water input applied at a constant rate for one hour. To derive the one hour UH for Rio
533 Behar catchment for each climate/SMB model, we used effective melt (M') and observed runoff
534 (Q) as input/output to derive the one hour UH transfer function (i.e., $Q = M' * UH$) using the
535 traditional optimization (20) as follows:

$$Q_{N \times 1} = M'_{N \times J} U_{J \times 1}$$

536 where $Q_n = \sum_{k=1}^{K} M'_k U_{n-k+1}$, N is total number of discharge measurements (N = 42), K is the
537 total number of hours of effective melt (K = 18), J is the number of hours in unit hydrograph (J
538 = N − M + 1). Therefore, the duration of the derived one hour UH is 25 hours. Examination of
539 the resultant one hour UH's shows that for one hour of M' across the Rio Behar catchment, the
540 associated time to peak (t_p) for that hour is 5.5-6.5 hours, depending on the model (**Table S3**).
541 The corresponding peak discharge (h_p) is 0.09 to 0.11 hr^{-1}, signifying that 9-11% of the input of
542 one hour of M' contributes to peak discharge. Because recession flow extends into the
543 following day, we used the aforementioned two complete diurnal cycles of separated
544 hydrographs to calculate these UH parameters t_p and h_p for both days for each model, and also
545 average them for presentation in **Table S3**.

546

547 **5. Extension of Synthetic Unit Hydrograph (SUH) model to 799 internally drained catchments**
548 **(IDCs)**

549 To isolate the influence of surface drainage pattern (i.e. IDCs) upon the timing and volume of
550 surface meltwater delivery to moulins across a larger area of the GrIS ablation zone, we apply a
551 simple "lumped" morphometric routing scheme, the Snyder Synthetic Unit Hydrograph (SUH)
552 model (21) and a Gamma function (22) to a previously published (7) broad-scale (13,563 km^2)
553 map of 799 remotely sensed IDCs (including Rio Behar catchment) have confirmed stream/river
554 networks. This requires fitting the observed UH calibrated at Rio Behar to other IDCs, taking
555 into account their differing shapes and areas. Note that this extension assumes that ice surface
556 properties were similar for all 799 IDCs as Rio Behar during the time of our field experiment. To
557 help justify this, we limit the time of our broader SUH application to the time of our field
558 experiment (21 July 2015).

59   5.1 Derivation of Synthetic Unit Hydrograph (SUH) model parameters for 799 IDCs:

560   A distinct advantage of the SUH method is its non-reliance on digital elevation models, thus

561   avoiding known challenges with the use of DEMs for modeling supraglacial hydrology, notably

562   selection of a user-specified parameter used to fill noise and/or true topographic depressions in

563   the ice surface DEM (23, 24), as well as breaching of headwater stream channels across

564   topographic ice divides (13). Rather than using DEMs, the SUH approach reduces the influence

565   of IDC morphometry to just three simple parameters, the total catchment area (A, in km^2), the

566   catchment main-stem stream length (L, in km), and an elongation proxy, calculated as the

567   distance from the catchment outlet (here, the terminal moulin) to the point on the main

568   channel nearest to the catchment centroid (L_c, in km). The Snyder SUH uses these watershed

569   metrics to estimate the aforementioned UH summary parameters t_p (time-to-peak, in hours)

570   and h_p (peak discharge, in hr^{-1}) for ungauged watersheds as $t_p = C_t(L_c)^{0.3}$, and $h_p = C_p/t_p$. Peak

571   discharge (Q_{pk}) at the catchment outlet (i.e. moulin) is $Q_{pk} = A*M'*h_p$. With unit conversion the

572   formula becomes $Q_{pk} = 0.28 *M' * h_p$, with Q_{pk} in m^3 s^{-1}, A in km^2, and M' in hr^{-1}. We used

573   ArcGIS to extract A, L, and L_c for 799 internally drained catchments previously mapped across

574   the southwest GrIS from a high-contrast 19 August 2013 panchromatic Landsat-8 image (7).

575   Note that C_p and C_t are dimensionless coefficients which we calibrate at Rio Behar catchment

576   using the field-calibrated UH, and L=13.8 km, and L_c =7.6 km from the Yang and Smith dataset.

577   A complete set of model-tuned estimates for C_p and C_t are presented in **Table S3**, with an

578   average value of C_p = 0.60 and C_t = 1.49. These ice-calibrated values fall within the normal

579   range of terrestrial values (0.56-0.69) for C_p and are lower than terrestrial values (1.8-2.2) for C_t

580   (21), however, at least two other studies report lower values of C_t in terrestrial watersheds, e.g.

581   0.3 to 0.7 (25), 0.4-2.4 (26), signifying that our low on-ice values do have precedent on land.

582   This lower value of C_t reduces t_p and increases h_p, indicating that Rio Behar catchment has a

583   flashy response to meltwater input relative to most terrestrial catchments.

584   The SUH coefficients C_p and C_t calibrated at Rio Behar catchment characterize the overall bulk

585   properties of water drainage efficiency off the ablating ice surface, and are held constant in the

586   present study. To examine the role of IDC-specific morphometry differences (i.e. A, L, and L_c)

587   on the timing and volume of runoff departing each of the 799 IDCs, we apply the SUH to them

588   using the simplest theoretical case (assuming a spatially uniform depth of 1 cm meltwater

589   runoff over a duration of 1 hour is allowed to drain from each IDC), and using climate/SMB

590   model output (M') to drive the SUH. For the simple case, we map the theoretical time to peak

591   (t_p) and peak moulin discharge (Q_{pk}) responses of a 1 hour uniform release of 1 cm of M' for all

592   799 IDCs (**Figure 4**). Note considerable uncertainties in small catchment boundaries at low

593   elevations (7) causing L_c values to be underestimated, making h_p values incorrectly large. This

594   explains why some small catchments at low elevations yield anomalously large peak discharge

595   (**Figure 4b,** also **Figure 5 and Figure S11**). Because a uniform melt depth is assumed

596 everywhere, **Figure 4** isolates the pure effect of catchment morphometry on t_p and Q_{pk}

597 independent of spatial and/or elevational gradients in meltwater production (i.e. surface mass

598 balance).

599 5.2 Generation of model-driven synthetic hydrographs for 799 IDCs:

600 To incorporate the impact of elevational and spatially varying gradients in melt production and

601 surface mass balance, we used climate/SMB model outputs of M' from MAR3.6, RACMO2.3,

602 and HIRHAM5 to drive the SUH model (**Figure 5, Figure S11**). Because point-SEB is a non-

603 gridded model driven by the KAN_M AWS, and MERRA-2 does not supply M), only MAR,

604 RACMO, and HIRHAM could deployed in this way.
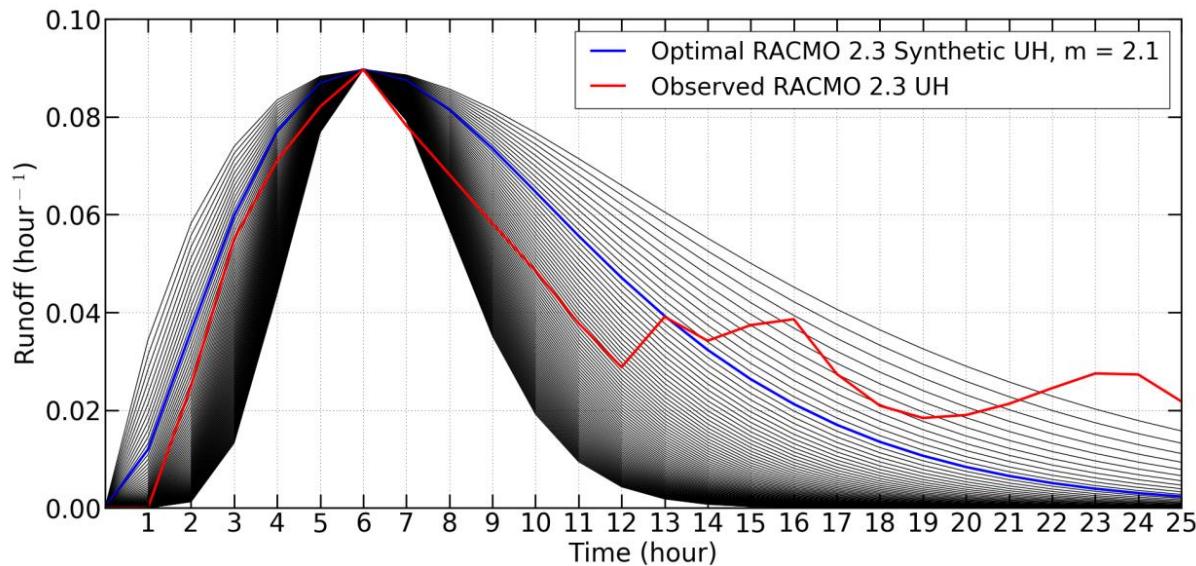

605

606 **Figure S6. Optimal model-specific synthetic unit hydrographs (SUHs) for Rio Behar catchment**

607 **may be approximated using Gamma functions, by varying the Gamma shape coefficient _m_ to**

608 **optimally fit each climate/SMB model using Nash-Sutcliffe model efficiency (see Equation 2).**

609 **For illustration, the observed RACMO 2.3 UH appears in red and its corresponding**

610 **approximated SUH in blue (shape coefficient m=2.1).**

611 To produce **Figure 4, Figure 5, and Figure S11** We built a SUH curve for each IDC based on

612 Snyder time-to-peak (t_p), peak discharge (h_p), and a Gamma function:

613
$$\frac{q}{h_p} = e^m \left[\frac{t}{t_p}\right] \left[e^{\left(-m\left(\frac{t}{t_p}\right)\right)}\right] \quad (1)$$

614 where q is the SUH discharge at time t, and m is the Gamma equation shape factor (22). The

615 parameter m controls the shape of Gamma function and must be determined empirically.

616 Recall that the SUH parameters t_p and h_p are different for each IDC, whereas C_p, C_t and m are

617 empirically calibrated from our Rio Behar field experiment and assumed to be constant in this
618 study. To determine model specific m values, we tested different m values using Nash-Sutcliffe
619 model efficiency (Equation (2)) for our model-specific Rio Behar Unit Hydrographs (**Figure S5b**),
620 to obtain the optimal value of m (e.g. for RACMO2.3, **Figure S6**). These model-optimized m
621 values are supplied in **Table S3**. To compute the theoretical peak moulin discharges presented
622 in **Figure 4b**, an optimal value of m of 2.5 was obtained from the 4-model average UH
623 (including SEB, thick red line **Figure S5b**).

$$E = 1 - \frac{\sum_{t=1}^{T}(Q_o^t - Q_m^t)}{\sum_{t=1}^{T}(Q_o^t - \overline{Q_o})} \qquad (2)$$

624

625 The SUH curve derived for each IDC estimates how meltwater production (more specifically, the
626 corrected meltwater production M') across each IDC is released to its terminal moulin over
627 time. Unlike the instantaneous area-integration method, climate/SMB model output is
628 distributed over time. For each IDC, convolution of SUH with hourly model outputs of M' yields
629 a discharge hydrograph at the terminal moulin.

630 We used the function $Q = M'$ * SUH to calculate discharges at IDC outlets, where M' is effective
631 melt, SUH is the Synthetic UH curve obtained above, and * is the convolution operator. As
632 such, input melt for each IDC is temporally distributed over time and the discharge hydrograph
633 computed for the IDC terminal moulin. Note that this approach assumes all runoff is routed to
634 the terminal moulin with no further losses of water (beyond any losses already captured by the
635 calculation $c \times M = M'$).

636 To compare the difference between SUH-routed runoff and instantaneous (non-routed)
637 approaches, we used the above procedure to model moulin discharges for all 799 IDCs at 1400
638 local western Greenland time on 21 July 2015 (a peak melt production time, **Figure 5**) and 10
639 hours later at 2400 local time on 21 July 2015 (**Figure S11**). The top row in each figure shows
640 climate/SMB output of M', the middle row instantaneous area-integrated runoff, and the
641 bottom row SUH-routed runoff. The timing of t_p (i.e. **Figure 4a**) remains unchanged so is not
642 repeated in these figures.   Row (a) in each figure shows presents the expected gradually
643 varying, ramped elevational gradient in M'. Row (b) of each figure shows instantaneous area-
644 integrated runoff (which assumes instant arrival of climatologically produced runoff at the
645 terminal moulin) for each IDC, as previously used to simulate meltwater injection into the ice
646 sheet (27, 28). Row (c) of each figure shows dramatic differences from the top two rows, with
647 **Figure 5c** (time 1400) moulin discharges lower than model output and nighttime moulin
648 discharges (time 2400 on 21 July 2015, **Figure S11c**) considerably higher.   Heterogeneous IDC
649 areas and shapes yield different t_p and Q_{pk}. For each IDC, time series of M', instantaneous area-
650 integrated R, and SUH-routed R was computed over the time period of 0000 to 2400 on 21 July
651 2015. From these time series the maximum (Q_{pk}) and minimum hourly runoff values were

652 extracted in ArcGIS software with Arcpy scripts, and difference to obtain daily diurnal difference
653 data (in units of discharge, $m^3 s^{-1}$). The time lag between peak climatological melt production
654 M' and Q_{pk} was also computed, and equaled zero for instantaneous area-integrated R, but up to
655 12 hours for SUH-routed R.   Summary scatterplots of these Q_{pk}, diurnal difference, and time-lag
656 data show lower peak discharges, less diurnal contrast, and non-trivial timing delays in SUH-
657 routed runoff and are presented in **Figure 6**.

658

659 <u>5.3 Retroactive validation of SUH using field data of McGrath *et al.* and Chandler *et al.*</u>

660 To validate our SUH routing model for other times and locations on the ice sheet, we applied it
661 retroactively to the previous field sites of McGrath *et al.* (6) and Chandler *et al.* (5) to determine
662 the ability of SUH to independently reproduce peak runoff time and lag time to peak t_p (i.e.,
663 time lag between peak melt generation across the catchment and peak discharge at the
664 terminal moulin) observed in their field measurements of supraglacial stream level. Owing to
665 known water leakage at these sites (e.g. ~48% runoff infiltration to crevasses (6)) and absence
666 of hourly climate/SMB model output prior to 2015, we did not assess Q_{pk} at these sites.
667 Hydrograph lag time to peak, however, is dominated by catchment morphometry (not absolute
668 magnitudes of melting and runoff), and so offers a useful independent test of the SUH approach
669 even in supraglacial catchments where runoff leakage to crevasses is known to occur.

670 **Table S4: Validation of SUH model with field data of McGrath *et al.* (6) and Chandler *et al.* (5)**

		Synthetic Unit Hydrograph (SUH)			Field-measured
		MAR 3.6	RACMO 2.3	HIRHAM 5	
McGrath et al. (2011)	Date	10 August 2015			3-17 August 2009
	Peak melt time	15:00	14:00	19:00	14:00 ± 231 min
	Peak runoff time	17:00	16:00	20:00	16:30-17:00
	Lag time to peak t_p	2 hours	2 hours	1 hour	2.8 ± 4.2 hours
Chandler et al. (2013)	Date	3 July 2015			29 June to 7 July 2011
	Peak melt time	13:00	13:00	19:00	-
	Peak runoff time	18:00	17:00	22:00	18:00-20:00
	Lag time to peak t_p	5 hours	4 hours	3 hours	-

671 The catchment studied by McGrath *et al.* (6) is in the Swiss Camp area, approximately 284 km
672 north of our field site. We obtained the same WorldView-1 image used by McGrath *et al.* (6)
673 (acquired on 15 July 2009, catalog ID: 1020010008AB4800) and followed McGrath *et al.* (6) to
674 map the catchment boundary. We then created the catchment centroid and manually
675 measured the main-stem stream length. The resultant catchment area (A_{idc}) is 1.1 km^2, main-
676 stem length (L) is 2.1 km, and the distance from the catchment outlet (i.e. terminal moulin) to
677 the point on the main channel nearest to the catchment centroid (L_c) is 1.2 km. We then

26

678 generated a SUH for this catchment as per Methods **Section 5.2**. Finally, the peak climate/SMB
679 melt time, peak runoff (discharge) time, and lag time-to-peak t_p were calculated. Depending on
680 choice of input climate/SMB model, the resultant SUH-derived t_p values range from $1 - 2$ hours
681 (**Table S4**), which is smaller but comparable to the field-based 2.8 ± 4.2 hours (6). Depending
682 on input model, peak runoff times ranged from 16:00 to 20:00, comparable to 16:30-17:00 in
683 the field measurements of McGrath et al.

684 Chandler *et al.* (5) also reported the peak river discharge time for an IDC (moulin site L41)
685 during 29 June to 7 July 2011. We used a WorldView-1 image acquired on 12 July 2011 (catalog
686 ID: 103001000CB46800) to delineate the IDC catchment boundary and mainstream as per
687 above. The resultant A_{idc} is 18.2 km^2, L is 6.9 km, and L_c is 3.6 km. The resultant SUH-derived
688 peak runoff time is 17:00-22:00, which matches well with the peak of 18:00-20:00 observed in
689 the field measurements of Chandler et al. (**Table S4**). In sum, despite using hourly climate
690 model output from 2015 to drive the SUH model, values of C_p, C_t and m calibrated at Rio Behar
691 catchment, and application to different locations and years, we find comparable values in the
692 timing of peak moulin runoff between these two previously published field studies and their
693 respective SUH-routed values as determined retroactively using their remotely sensed fluvial
694 catchments.

695

696 **Methods 6. Regional climate/SMB model descriptions and data analysis**

697 Hourly simulations of GrIS meltwater production M and runoff R over the study period were
698 generated using the HIRHAM5, MAR3.6, RACMO2.3, MERRA-2, and point SEB models with
699 descriptions as follows. Note that there is no overland surface catchment routing scheme
700 present in any of the model simulations in this study. MAR has a built-in routing delay intended
701 to represent meltwater passage through the ice sheet to its edge; this delay is left in place for
702 **Figure 3** and **Figure 7a** but is eliminated in **Table S5**. HIRHAM5 does introduce a delay for
703 runoff from snow (slush), but melting over bare ice converts to immediate runoff.   Most
704 current models do not consider penetration of shortwave radiation and associated subsurface
705 melting (29), so this option was turned off in point SEB, to maintain consistency across models.

706 6.1 HIRHAM5:

707 The HIRHAM5 regional climate model (30) is developed by the Danish Meteorological Institute
708 and the Potsdam Research Unit of the Alfred Wegener Institute Foundation for Polar and
709 Marine Research. It has a native horizontal spatial resolution of 5.5×5.5 km ($0.05°$x$0.05°$ on a
710 rotated pole grid (31)) with 8 native-resolution grid cells intersecting Rio Behar catchment. It
711 combines the dynamics of the <u>HIR</u>LAM weather forecast model (32, 33) with the physical

712 parameterization schemes of the ECHAM climate model (34). Six hourly inputs of horizontal
713 wind vectors, temperature, and specific humidity from the ERA-Interim reanalysis dataset (35)
714 are supplied at the domain boundaries at all atmospheric levels to compute the atmospheric
715 circulation within the domain at 90 s time steps. The resulting surface fluxes of energy
716 (turbulent and downward radiative) and mass (snow, rain, evaporation, and sublimation) are
717 used to drive an offline snow/ice subsurface scheme which provides SMB, runoff and refreezing
718 rates (36). A number of updates have been made to the subsurface scheme compared to the
719 version used by Langen *et al.* (36). The surface energy budget calculation incorporates daily
720 observed MODIS MOD10A1 surface albedo de-noised after Box *et al.* (37). Snow undergoes
721 temperature and pressure dependent densification (38).

722 As illustrated by Lucas-Picher *et al.* (31), ice sheet accumulation is quite accurately represented
723 with biases in the south generally smaller than 10% compared to ice core–derived accumulation
724 rates (31). In the Nuuk area, runoff was found to be underestimated by 10-20%, mainly due to
725 too high albedo in the lower ablation zone (36). With the MODIS-derived albedos employed
726 here, this effect is expected to be limited. HIRHAM5 is run at higher vertical resolution than in
727 Langen *et al.* (36), thus employing 25 layers with a total water equivalent depth of 70 m.
728 HIRHAM5 does not employ any nudging inside the model domain, and the driving atmospheric
729 fields from ERA-Interim are thus only felt on the domain's lateral boundaries. The aerodynamic
730 roughness length for momentum (z_0) is set to a constant value of $z_0 = 1$ mm for both snow and
731 bare glacier ice.

732 When snow/firn is present and a layer bulk density exceeds the pore close off density, water
733 percolating from above is treated as a slush layer that runs off with a time scale depending on
734 surface slope (39, 40). During bare ice conditions, however, any melt that occurs is
735 immediately converted to runoff with no delay. For this particular study, hourly output was
736 supplied from January 1 to August 31 2015, following a spin-up period of 70 years. Rain is
737 parameterized in the model but did not occur at the study site during the study period.
738 HIRHAM5 assumes all energy fluxes to balance at the surface skin layer with no allowance of
739 shortwave radiation penetration and associated subsurface melting.

740 6.2 MAR3.6:

741 The Modèle Atmosphérique Régionale (MAR) is a modular atmospheric model that uses the
742 sigma-vertical coordinate to simulate airflow over complex terrain (41, 42) and the Soil Ice
743 Snow Vegetation Atmosphere Transfer scheme (SISVAT) (43, 44) as the surface model. It has a
744 native horizontal spatial resolution of 20 × 20 km with 2 native-resolution grid cells intersecting
745 Rio Behar catchment. The snow–ice part of SISVAT, based on the CEN (Centre d'Etudes de la
746 Neige) snow model called CROCUS (45), which calculates albedo for snow and ice as a function
747 of snow grain properties, which in turn are dependent on energy and mass fluxes within the

748   snowpack. CROCUS is a one dimensional multilayered energy balance model that determines
749   the exchanges between the sea ice, the ice sheet surface, the snow-covered tundra, and the
750   atmosphere. It allows meltwater refreezing and snow metamorphosis, influencing the
751   transformation of snow to ice and the surface albedo using the CROCUS formulations(45, 46).
752   For snowpack having surface density > 550 kg m^{-3} (representing the maximum density of
753   pure snow), the minimum allowed albedo is calculated linearly as a smooth function
754   between pure snow albedo (0.7) and clean ice (0.55) (47).

755   The lateral and lower boundary conditions are prescribed from meteorological fields modelled
756   by the global European Centre for Medium-Range Weather Forecasts (ECMWF) Interim
757   Reanalysis (ERA-Interim, http://www.ecmwf.int/en/research/ climate-reanalysis/era-interim).
758   Sea-surface temperature and sea-ice cover are also prescribed in the model using the same
759   reanalysis data. The atmospheric model within MAR interacts with the CROCUS model, which
760   provides the state of the snowpack and associated quantities (e.g. albedo, grain size). No
761   nudging (assimilation of AWS meteorological data to improve model performance) or interactive
762   nesting was used in any of the experiments. This is not done with MAR, which uses only reanalysis
763   data as input to its atmospheric model. The aerodynamic roughness length for momentum is set
764   to a constant value of z_0 = 0.1 mm for dry snow. For melting snow or ice z_0 is a function of
765   density and varies between 1 and 3 mm for both snow and bare glacier ice. An optional MAR
766   representation for shortwave radiation penetration and associated subsurface melting into
767   bare ice has been developed but to maintain consistency with other models was not used here.

768   The Greenland topography used for our simulations was derived from the high-resolution (5
769   km) digital elevation model from radar altimetry (48, 49), and the ice sheet mask is based on
770   the Greenland land surface classification mask from Jason Box (http://bprc.osu.edu/wiki/Jason
771   Box Datasets) using MODIS calibrated radiances imagery.

772   MAR is the only climate/SMB model integrating a runoff delay function to retard bare-ice
773   surface runoff over time. This delay function was proposed by Zuo and Oerlemans (40) based
774   on the idea that surface meltwater probably reaches the supraglacial rivers quicker when the
775   general surface slope is larger. Lefebre *et al.* (39) updated the coefficients of this delay function
776   to route meltwater more quickly. This MAR delay function describes the time lag from surface
777   meltwater production to its drainage through the ice sheet to its edge.

778   <u>6.3 RACMO2.3:</u>

779   The RACMO2 regional climate model uses the atmospheric dynamics module from the High
780   Resolution Limited Area Model (HIRLAM) and adopts the physics package of the European
781   Centre for Medium-range Weather Forecasts Integrated Forecast System (ECMWF-IFS)(32, 33).
782   It has a native horizontal spatial resolution of 11 × 11 km with 3 native-resolution grid cells

783   intersecting Rio Behar catchment. For a detailed description of the basic version of Regional

784   Atmospheric Climate Model (RACMO2) the reader is referred to Van Meijgaard *et al.*(50). A

785   polar version of RACMO2 has been developed by the Institute for Marine and

786   Atmospheric Research (IMAU), Utrecht University, and is especially adapted for use over

787   ice sheets and glaciated regions. It is interactively coupled to a multilayer (N_{max} = 100), 1-

788   dimensional snow model, accounting for meltwater percolation, refreezing and runoff (51); a

789   snow albedo scheme with prognostic snow grain size (52, 53) and a drifting snow module,

790   simulating snow erosion and drifting snow contribution to sublimation (54). For ice albedo an

791   11 km version of the 500 m Moderate-resolution Imaging Spectroradiometer (MODIS) 16-days

792   Albedo product (MCD43A3) is used. Bare ice albedo is estimated as the averaged 5% lowest

793   surface albedo measurements for the period 2001- 2010 (55). The firn layer is initialized using

794   3D temperature and density fields from previous runs with a dedicated firn model (52, 53).

795   Every six hours, RACMO2 is forced at the lateral boundaries by ERA-Interim reanalysis data

796   (1958-2015). Recently, RACMO2 has been updated to version 2.3 (56) and leading to

797   improved representation of GrIS SMB (55). The present study uses this latest version

798   RACMO2.3. The model has proven to realistically simulate SMB and climate of the GrIS, as well

799   as the extent of the perennial firn aquifer in southeast Greenland (57). SMB gradients are

800   well captured, but accumulation in the interior ice sheet appears underestimated by 5-10%

801   (integrated value). Other perceived weaknesses of the model are the assumptions of

802   temporally constant ice albedo, and instantaneous runoff. RACMO2.3 does not employ any

803   nudging inside the model domain, and the driving atmospheric fields from ERA-Interim are thus

804   only felt on the domain's lateral boundaries. The model assumes fixed values for roughness

805   length of z_0 = 1 mm over snow and 5 mm over bare ice.

806   RACMO2.3 assumes all energy fluxes to balance at the surface skin layer with no allowance of

807   shortwave radiation penetration and associated subsurface melting. No time delay is

808   introduced between melt generation and runoff.

809   <u>6.4 MERRA-2:</u>

810   MERRA-2 is a global atmospheric reanalysis produced by the NASA Global Modeling and

811   Assimilation Office (GMAO) for the satellite observing era from 1980 until the present at a grid

812   spacing of ½° latitude by ⅝° longitude and 72 hybrid-eta levels from the surface to 0.01 hPa

813   (58). Over our field area it therefore has a native horizontal spatial resolution of 56 x 28 km

814   with 1 native-resolution grid cell intersecting Rio Behar catchment. MERRA-2 serves as an

815   update on the previous MERRA product (59) by incorporating radiance data from more recent

816   satellites including NOAA-19, MetOp-A and -B, and the Suomi National Polar-orbiting

817   Partnership (Suomi-NPP). The background model is the Goddard Earth Observing System

818   model, version 5 (GEOS-5). The model uses a finite-volume dynamical core (60) that is

819   integrated with various physics packages. These physical packages incorporate several

820   improvements, which are described in Molod *et al.* (61). Additionally, MERRA-2 incorporates

821   several new features including an interactive aerosol analysis, a scheme to conserve globally-

822   averaged atmospheric mass and moisture (62), and the use of a cubed sphere grid for

823   computations. The representation of glaciated land surfaces has been updated as described

824   and evaluated (63). The model represents energy conduction properties of the upper 15 m of

825   glacial ice, and energy and hydrologic properties of an overlying, variable snow cover.

826   Snow hydrology follows a modified version of the Stieglitz model, which provides an explicit

827   representation of snow densification, meltwater runoff, percolation, refreezing, and surface

828   albedo (64, 65). Over land ice, the snow pack is vertically discretized into fifteen layers, which

829   are demarcated by fractions of the total snow depth. Firn of density greater than 500 kg m^{-2} is

830   not explicitly represented; this provides an approximate upper limit on the total depth of the

831   snow pack. Snow cover is also allowed to be fractional. A prognostic surface albedo is based on

832   Greuell and Konzelmann (66). Bare ice albedo is set to 0.58. As described in Lynch-Stieglitz (65),

833   meltwater is generated when the heat content of a snow layer exceeds the minimum necessary

834   for the layer to remain entirely frozen (65). A liquid water holding capacity is defined for each

835   snow layer (67). Meltwater exceeding the layer holding capacity is transferred to the next

836   lowest layer. Liquid water leaving the lowest model layer is instantaneously designated as

837   runoff. The fractional bare ice cover may also generate runoff based on the excess melt energy

838   from the surface energy budget. No delay is introduced between melt production and runoff.

839   For efficiency, MERRA-2 was integrated in four processing streams: 1980-1991, 1992-2000,

840   2001-2010, and 2011 to the present. A one-year overlap for each stream was incorporated to

841   avoid temporal discontinuities in the transitions, particularly in land surface variables. It was

842   recognized that the annual temperature wave would not reach lower levels of the surface

843   representation over glaciated land. As a result, glaciated land variables are restricted to those

844   describing albedo, fractional snow cover, and runoff. Like RACMO2.3, MERRA-2 assumes fixed

845   values for roughness length of $z_0 = 1$ mm over snow and 5 mm over bare ice. RACMO2.3

846   assumes all energy fluxes to balance at the surface skin layer with no allowance of shortwave

847   radiation penetration and associated subsurface melting. Meltwater production *M* is not

848   supplied by MERRA-2, thus precluding consideration of MERRA-2 from many parts of this

849   analysis.

850   <u>6.5 Point SEB:</u>

851   The point-based SEB model (68) calculates SMB using in-situ measurements from the KAN_M

852   automatic weather station (67.0667 N, 48.8327 W, 1270 m a.s.l.; **Figure 1**). Being a point-based

853   model, the outputs driven by the KAN_M station were simply extrapolated to the rest of Rio

854   Behar catchment. The model's inability to calculate spatial variability within the domain is a

855 trade-off for accurate, local forcing at hourly time steps. Measurements of absorbed shortwave
856 (solar) radiation and downward longwave (terrestrial) radiation feed into the model. The
857 turbulent heat fluxes are calculated similarly to those produced by the RCMs, using near-
858 surface gradients in temperature, wind speed and specific humidity to approximate vertical
859 transport of sensible and latent heat. The sub-surface thermal calculations are performed to a
860 depth of 20 m with 0.2 m spacing, which is initialized using thermistor string measurements of
861 ice temperature, and assuming a constant ice density and thermal conductivity for this ablation
862 area site in summer. The energy flux calculations of upward longwave radiation, sensible heat,
863 latent heat and sub-surface heat make use of surface temperature; the model iteratively
864 determines the surface temperature for which all surface energy fluxes are in balance. If
865 surface temperature is capped by a 0°C melting surface, the surplus energy determines the
866 melt rate. The calculated meltwater at the ice sheet surface can theoretically refreeze in sub-
867 surface model firn layers, yet for this study's location and observational period no snow or firn
868 was present at the AWS. Therefore modeled surface meltwater runoff equals the meltwater
869 production, with minor compensation for condensation and evaporation. Rain is
870 parameterized in the model but did not occur at the study site during the study period. For the
871 study region and season the 1-dimensional SEB model assumes an aerodynamic roughness
872 length of 0.1 mm following Smeets and van den Broeke (69).

873 We used KAN_M measurements of incoming and outgoing shortwave radiation to determine
874 how much energy was absorbed at the surface, so did not assume a pre-defined albedo. The
875 SEB model does provide the possibility to allow shortwave radiation to penetrate the surface
876 following Beer's law and generate subsurface melting. However, we did not use that option in
877 our model calculations for several reasons: 1) to enable consistency with the other models that
878 don't allow penetration; 2) a necessary extinction coefficient for the high dust/high algae
879 content ice at KAN_M is currently unknown; 3) a necessary runoff threshold for meltwater
880 generated within the ice matrix is currently unknown. In sum, the SEB radiation penetration
881 scheme requires further research and development before implementation.

882 6.6 Reprojection of all model outputs to a common resolution and grid:

883 To improve comparison among these models, their outputs of meltwater production *M* and
884 runoff *R* were reprojected to a common 5 km posting and map projection (i.e. to that of MAR)
885 using a 'drop in a box' (nearest neighbor) resampling. This method was chosen over use of an
886 interpolation scheme, to preserve the native model output with in situ field measurements.
887 Thus the native grid cell resolutions remain visible in **Figure 5a** and **Figure S11** despite finer-
888 scale resampling to 5 km.

889 The common projection chosen is Polar Stereographic (70) based on the WGS84 ellipsoid, with
890 true scale at $71^{°}$ S and posting of 5 km. The map reference latitude was set at $90^{°}$ N and

891 reference longitude at 39° W. Map origins were adjusted per MAR outputs. For area and
892 volume calculations, the Lambert Azimuthal Equal Area projection was used (70). Latitude and
893 longitude values for the different datasets were converted to map co-ordinates. Finally, the
894 reprojected variables were interpolated onto the 5 km grid (48, 49), using nearest-neighbor
895 interpolation. Note that this sampling is generally finer than the native resolutions of most
896 models, allowing smoother interpolation across the watershed.

897 Two exceptions to the above processing stream are the MERRA-2 and point SEB models.
898 Because MERRA-2 does not provide M as a model output, assessments of melt production,
899 runoff ratios, and SUH parameters are not possible for this model. Because SEB is a point-
900 based, non-gridded model driven by an AWS, outputs of M and R were applied uniformly across
901 the Behar catchment surface area without adjustment, owing to identical elevation and close
902 proximity of the KAN_M AWS station to Behar catchment.

903 6.7 Comparison of model outputs with field observations from Rio Behar catchment:

904 **Table S5: Climate/SMB model and field measurements of M and R**

	Melt (mm)	Runoff (mm)	Difference (mm)
MAR 3.6*	51.1	41.7	9.4** (-0.4)
Point SEB	46.7	47.8	-1.1
RACMO 2.3	51.1	51.7	-0.6
HIRHAM 5	67.6	61.0	6.6
MERRA 2	N/A	30.4	N/A
Observation (ablation stakes, ADCP)*	19.0-26.9	31.4	-12.4 - -4.5
Observation (KAN-M, ADCP)	16.2-23.0		-15.2 - 8.4
Overestimation/underestimation by models	+73.6% to +317.3%	-3.4% to +94.2%	75.6% - 246.7%

905 *MAR runoff R and observed ADCP discharge are lagged by 5 hours to secure peak-to-peak matching

906 **Value 9.4 is unrealistic due to delay-to-edge R smoothing unique to MAR. The more appropriate
907 value, calculated directly from raw MAR SMB variables is -0.4 mm.

908 For the four gridded climate/SMB models (HIRHAM5, RACMO2.3, MAR3.6, MERRA-2) hourly
909 model outputs of melt M and runoff R were computed within Behar catchment using Python
910 and ArcGIS (13, 71). This procedure entails clipping the 5 km × 5 km model grid cells with the
911 catchment polygon boundary, weighting the runoff/melt values of each grid cell by the percent
912 area contained within the catchment, and summing their corresponding runoff/melt values to
913 compute the total runoff/melt (m^3/s) inside the catchment (13). Note that for presentation of
914 climate/SMB model output of R, there is no physical difference between use of units of
915 discharge Q (m^3/s) and runoff depth (mm/hour) they are interchangeable via unit conversion

916   using the mapped catchment area (km^2). Presentation of modeled runoff R in units of

917   discharge (m^3 s^{-1}) is derived by multiplying model output of R (in units of mm hr^{-1}) times our

918   minimum, mid- and upper catchment areas (km^2), to obtain units of discharge (m^3 s^{-1}) suitable

919   for direct comparison with in situ ADCP measurements. The upper and lower uncertainty

920   bounds on modeled runoff (in units of Q) of **Figure 3** and **Figure 7a** thus reflect model

921   uncertainty due solely to catchment area uncertainty.

922   For a 37.5 hour melt-production period with overlapping ADCP discharge and ablation stake

923   measurements (over the overlapping period July 21, 11:00 am to July 23, 00:30 am local

924   Greenland time, see **Table S5**) the various models overestimated runoff by -3% to +94%, due to

925   the combined effects of overestimated melt production (~8% to 57% overestimation) and

926   underestimated water retention/refreezing processes (~44% to 109% underestimation). For

927   these runs, the difference between M and R in the models is driven by modeled meltwater

928   retention and/or refreezing processes, except for a brief, minor rainfall event modelled by SEB,

929   MAR 3.6, and RACMO 2.3 at approximately 22:00 on 22 July, which added to the runoff

930   calculation thus increasing it slightly over R and small negative differences in **Table S5**. Note

931   that runoff = melt - refreezing + rain + condensation - evaporation, so melt and runoff are not

932   expected to be identical. In particular, runoff may exceed melt due to rain and/or

933   condensation.

934   The seemingly large M – R difference for MAR3.6 (9.4 mm in **Table S5**) is in fact an artifact of

935   this particular model's aforementioned delay-to-edge smoothing of standard model output of R

936   (this feature also greatly smooths the MAR temporal runoff signal as seen in **Figure 3 and**

937   **Figure 7a**). A clearer view of how MAR really works is provided by calculating R directly from

938   the raw model data (R = melt + rain - refreezing - evaporation - sublimation) as per **Figure S7**.

939   This figure clearly confirms that $M \sim R$ in the MAR3.6 simulations for Behar catchment (and

940   indeed are virtually identical), with negligible retention/refreezing of runoff. The slightly

941   negative M-R difference (-0.4 mm, **Table S5**) results from a modelled trace rainfall event.

942
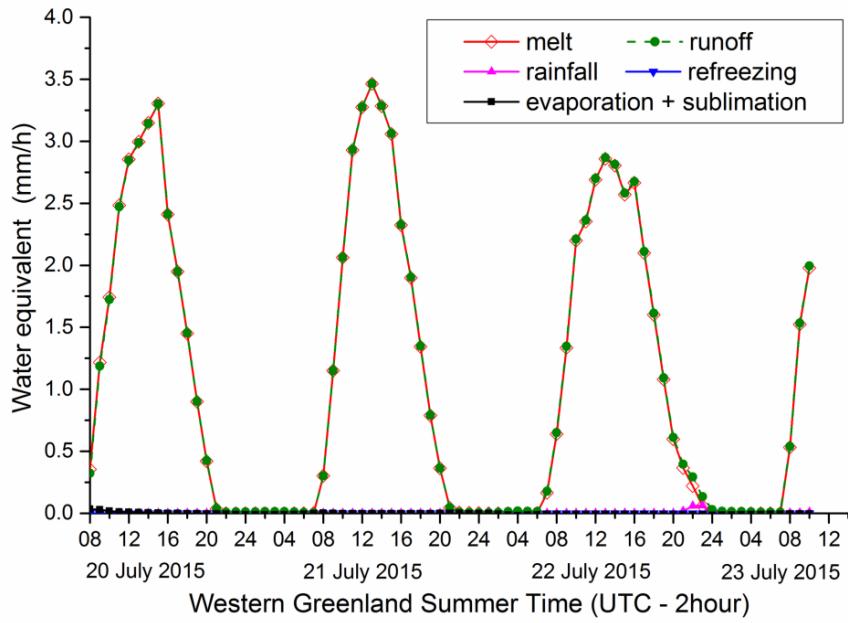
Figure S7: Raw SMB components from the MAR 3.6 model over the experiment period show zero or minimal refreezing, sublimation, evaporation, and rainfall, hence _R_ ~ _M_ with negligible meltwater retention simulated for the ice sheet surface. Note that direct calculation of runoff _R_ from these raw MAR variables (i.e. runoff = melt + rain - refreezing - evaporation - sublimation) eliminates the delay-to-edge smoothing applied to MAR _R_ standard output, providing the more realistic M-R difference (-0.4 mm, with negative value due to a trace rainfall event) shown in Table S5.

Because MERRA-2 does not supply outputs of melt _M_, we cannot directly confirm that its lower estimates of runoff _R_ are due to better approximation of _M_ or some other driver. The three regional climate models HIRHAM5, RACMO2.3, MAR3.6 are forced with ERA-Interim reanalysis data, for example, whereas MERRA-2 performs its own reanalysis. However, our comparison of ERA-Interim and MERRA-2 geopotential height fields over the 20-23 July 2015 study period shows that two reanalysis datasets are virtually identical, with little difference between them during our field experiment (**Figure S8**). Furthermore, Point SEB was driven by in situ meteorological observation from the KAN_M AWS, not reanalysis, yet shows similar overestimation in ice surface lowering and runoff as the other models (Figure 3, Figure 7). Another possible difference with MERRA-2 regards snow availability: In the MERRA-2 configuration, a strong delineation occurs between fresh snow and bare ice, as opposed to the other models that have explicit representation of firn. Finally, because MERRA-2 computes a prognostic surface albedo (66) yielding a bare ice albedo of 0.58, its albedo is higher than the other models, thus reducing the amount incoming shortwave radiation converted to melting and runoff.

964

(a)

(b)

965

(c)
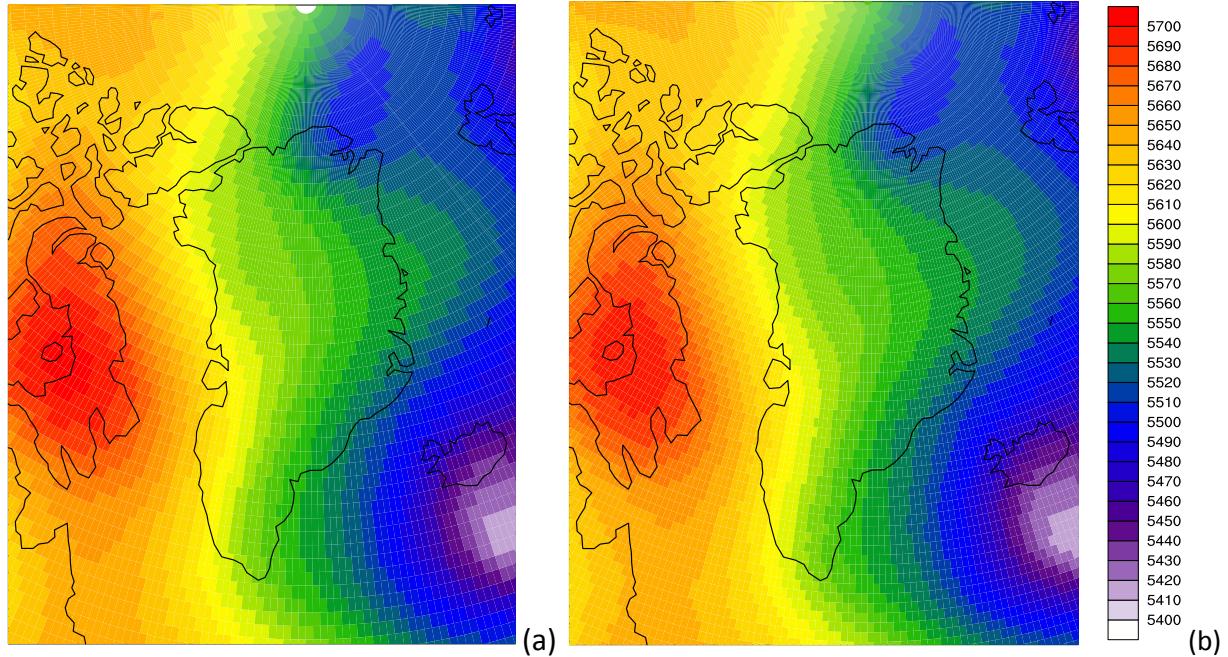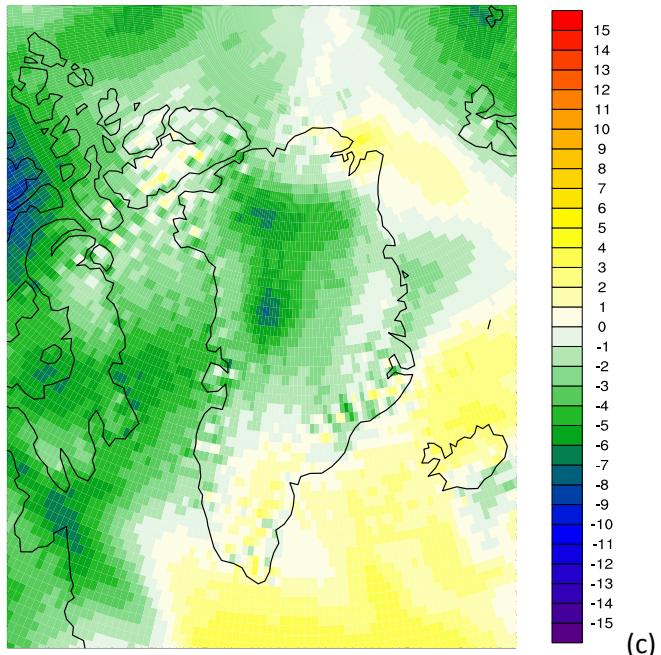
966 **Figure S8. Instantaneous 500 hPa geopotential height fields (values in meters) averaged for 0, 6, 12,**

967 **and 18Z during 20-23 July 2015 for: (a) ERA-Interim; (b) MERRA-2; the difference (c) of MERRA-2 minus**

968 **ERA-Interim. To generate (c) ERA-Interim data were interpolated to the MERRA-2 grid using spherical**

969 **harmonics. Representations of atmospheric dynamics during the study period are virtually identical in**

970 **both reanalysis datasets, with near-zero differences over southwestern Greenland.**

971

972   <u>6.8 Comparison of RACMO2.3 albedo and surface energy balance with AWS measurements</u>

973   To determine if discrepancies in modeled versus observed surface energy balance might explain
974   the observed discrepancies between modeled versus observed ablation and runoff, we
975   obtained hourly measurements of energy balance components collected by the KAN_M AWS
976   during our field experiment. These data were compared with hourly surface energy balance
977   outputs from RACMO2.3, as a representative of the four climate/SMB models that
978   overestimated surface runoff during the field experiment and for which detailed surface energy
979   balance outputs are available (**Figure S9**).

980   The RACMO2.3 albedo of 0.49 during the study period was close to AWS measurements (~0.45
981   to 0.55 over the entire period; ~0.46-0.50 during peak radiation hours, **Figure S9, panel 2**). We
982   also obtained MODIS satellite albedo retrievals (MYD10A1 daily product) and found that the
983   mean remotely sensed albedo across Rio Behar catchment was 0.43 on 20 July and 0.41 on 21
984   July - again, not far off from the modeled assumption of 0.49. This somewhat lower albedo
985   from MODIS may help to explain the higher melt and runoff estimates from HIRHAM5 relative
986   to the other models (recall HIRHAM is the only model to use MODIS satellite albedo retrievals),
987   but is of the wrong sign to explain the observed RACMO2.3 runoff overestimation. Together
988   with the good agreement between in situ and RACMO2.3 albedo, we therefore conclude that
989   underestimation of albedo cannot explain the observed overestimation of runoff.

990   To determine if discrepancies in modeled versus observed radiation effects of clouds might
991   explain the observed discrepancies between modeled versus observed ablation and runoff, we
992   compared RACMO2.3 shortwave and longwave radiation with measurements from the KAN_M
993   AWS. From the AWS longwave data and our own field notes, clouds moved into the study area
994   around 20:00 of Day 2 of the field experiment and persisted throughout Day 3. The arrival of
995   these clouds is missed for approximately 12 hours by RACMO2.3, as evidenced by a sustained
996   increase in longwave down (LWD) observed at KAN_M that is not immediately simulated by the
997   model (**Figure S9, panel 3**). The delayed detection of these clouds yields a small model
998   overestimation of shortwave down radiation (SWD), underestimation of LWD, and a small
999   overestimation of net SW radiation and net total radiation on Day 3 of the field experiment only
1000   (**Figure S9, panel 5**). This small model overestimation of net radiation did not occur on sunny
1001   Days 1 and 2, when climate/SMB models also overestimated observed ice surface lowering and
1002   runoff despite good simulation of net LW radiation and slight <u>underestimation</u> of net SW and
1003   net radiation. We therefore conclude that the delayed detection of clouds may have
1004   contributed slightly to the observed model overestimation of ice sheet ablation and runoff on
1005   Day 3, but not during the other two days of our field experiment when model overestimations
1006   also occurred.
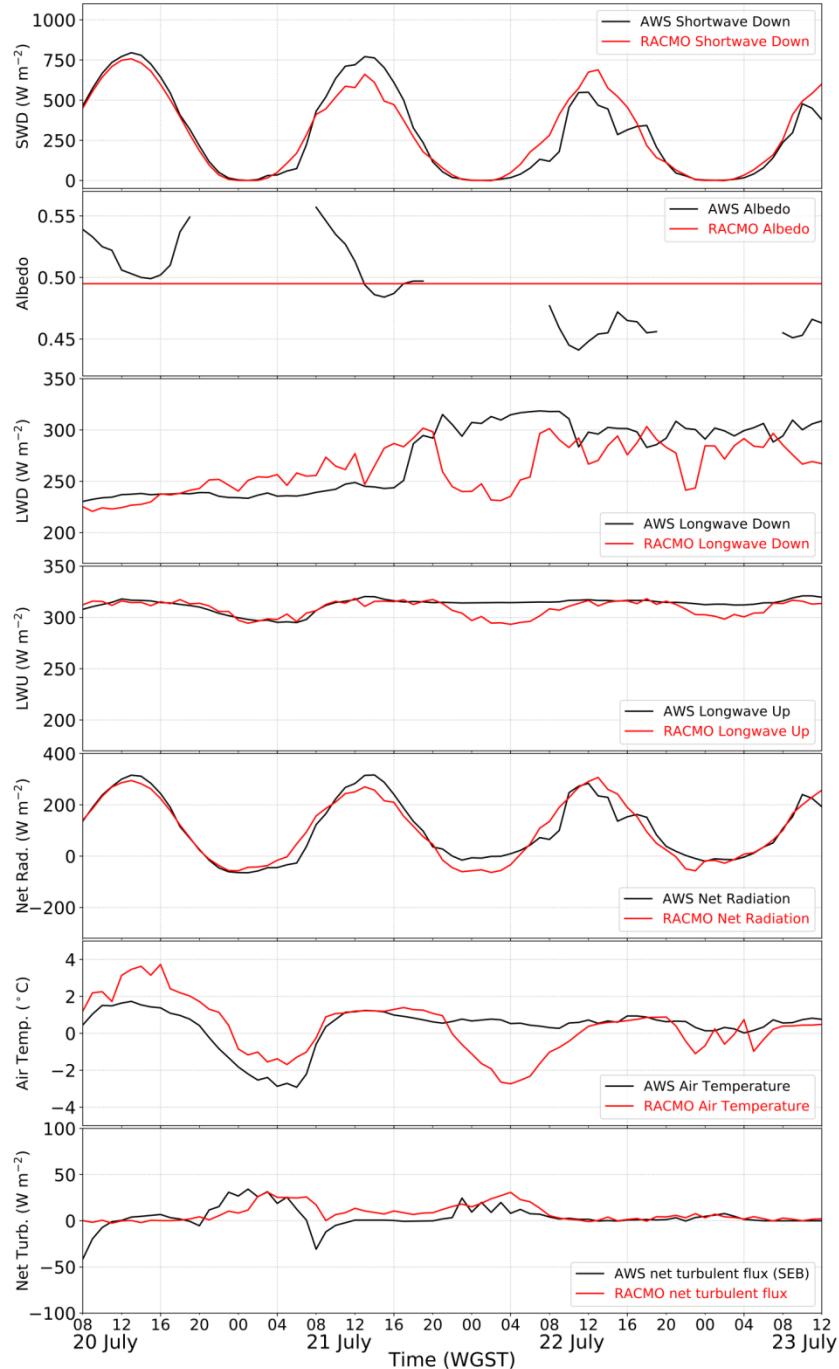
1007

Figure S9: Comparison of RACMO2.3 and KAN_M AWS surface radiation and energy balance

components during the 20-23 July 2015 Rio Behar field experiment. From top to bottom, black lines

show in situ measurements (shortwave down, albedo, longwave down, longwave up, net radiation,

surface air temperature) and net turbulent heat flux (calculated by the SEB model forced by KAN_M

AWS measurements). In general, RACMO2.3 reproduces the in situ surface energy balance quite well,

with small discrepancies in longwave down and net radiation insufficient to explain RACMO2.3

overestimation of observed ice surface lowering and runoff.

1015 Similarly, a generally good agreement between RACMO2.3 and AWS-derived net turbulent heat
1016 flux (**Figure S9, panel 7**), together with the small magnitude of this flux (less than ~25 W/m^2)
1017 suggest sensible and latent heat fluxes cannot explain the observed phenomena.

1018

1019 **Supplementary Discussion I: Weathering crust hypothesis for model overestimation of runoff**
1020 **ice surface lowering and supraglacial river runoff**

1021 For this particular field experiment, HIRHAM 5, MAR 3.6, RACMO 2.3, and SEB outputs of Rio
1022 Behar catchment R substantially overestimated our ADCP field measurements of catchment
1023 discharge (**Figure 7a**). If these models' outputs of M are converted to units of equivalent ice
1024 thickness they also substantially overestimated ice surface ablation (**Figure 7b**), even if a low
1025 near-surface ice density measured at our site (0.688 g cm^{-3})(3) is used for unit conversion
1026 instead of the density of solid ice (0.918 g cm^{-3}) (both density assumptions are presented in
1027 **Figure 7b**). Upon first examination, this suggests that climate/SMB models overestimated
1028 meltwater production M and hence R.

1029 However, an alternate, better-supported hypothesis is that the models estimated M correctly
1030 (or more precisely, the amount of energy allocated to M) but the ablation stakes and KAN_M
1031 surface lowering data underestimated it. A known process for this is shortwave radiation
1032 penetration and subsurface melting of bare ice (29), creating a porous, low density, ice matrix-
1033 supported "weathering crust" (72, 73). Because energy is expended to melt ice beneath the
1034 surface, but an ice matrix remains intact, this can produce less surface lowering than would
1035 occur from surface melting alone (3, 74, 75). Water saturated weathering crust was observed
1036 in abundance in our Rio Behar base camp in 2015, and again in 2016 when its depth exceeded
1037 1.1 m, the maximum length it could be cored (3) (**Figure S10**). Subsurface melting was not
1038 considered in the climate/SMB model simulations. If so, it is possible that our estimates of melt
1039 M derived from ablation stakes and at the KAN_M station may underestimate true melt
1040 production at the site.

1041 A separate bit of evidence supporting this comes from comparison of lagged ADCP discharge
1042 and ablation stake measurements for a 37.5 hour overlap period (**Table S5**), which suggest a
1043 seeming *surplus* of runoff (31.4 mm vs 16.2-26.9 mm). A surplus would signify $R > M$, which is
1044 nonsensical from an energy balance perspective. Again, it's hard to take this comparison too
1045 far because the ablation measurements are highly local, whereas our ADCP discharges integrate
1046 runoff over the entire catchment, but the data cannot rule out the possibility that the surface
1047 lowering data from our ablation stakes and the KAN_M station do not fully reflect the net
1048 surface energy balance that occurred during our field experiment. If so, the observed mismatch
1049 between modeled and measured R was likely caused by meltwater delay, retention, and/or

refreezing within the weathering crust itself. Weathering crust delays meltwater from reaching supraglacial channels (73, 76, 77) and is known to store and possibly refreeze meltwater at our field site (3). It is often accompanied with surface expression of water-filled cryoconite holes, which were ubiquitous around our base camp and the field area more generally, as observed visually from helicopter transit flights. Any refreezing of this meltwater, which we observed nightly in cryoconite holes, requires re-melting thus consuming a commensurate fraction of M the following day (or the following week, or the following year - the residency time of meltwater found in Rio Behar weathering crust is currently unknown). Small discrepancies in model-calculated versus measured ablation rates on the Haut Glacier d'Arolla, Switzerland, for example, likely result from the refreezing of surface water at night, its re-melting the next morning, and subsurface melting during the afternoon (78). In a model that does not simulate this process but otherwise correctly quantifies surface melt energy, both ice surface lowering and runoff would be overestimated.
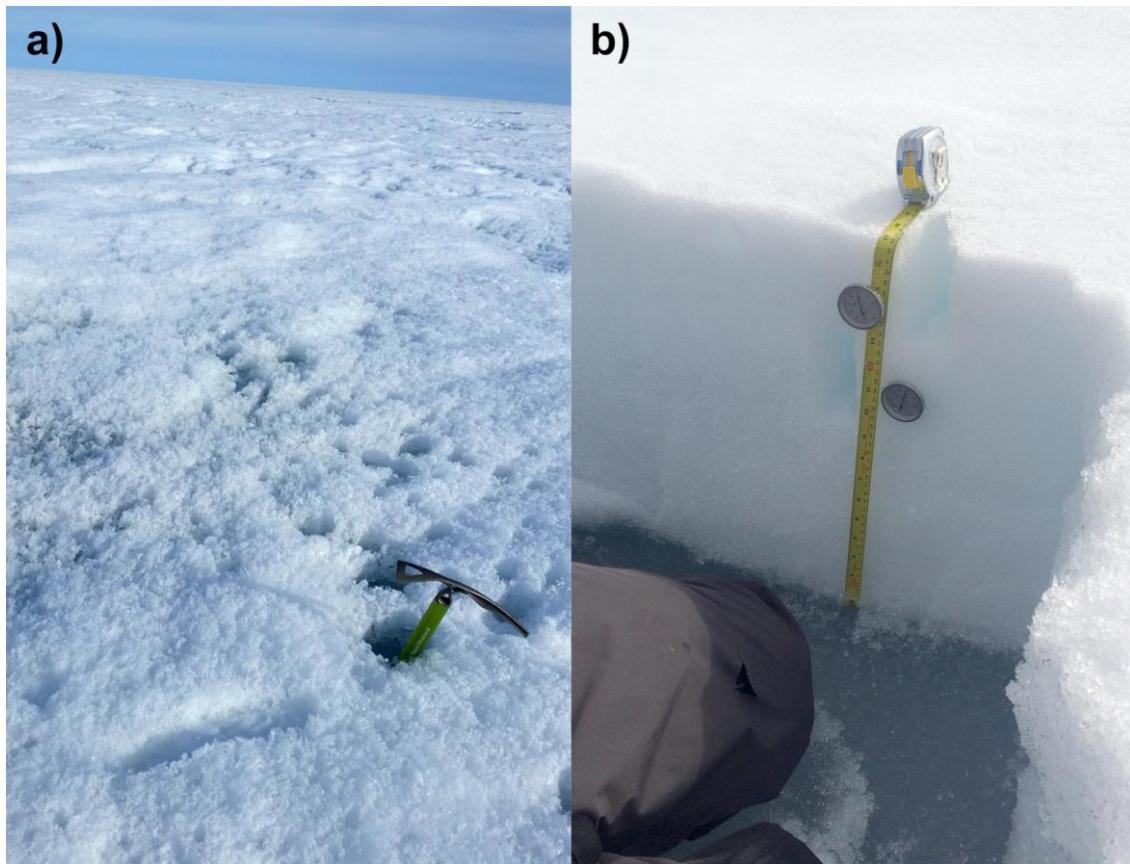

Figure S10: (a) The exposed, bare-ice ablation surface at Rio Behar base camp is characterized by weathering crust, a porous, water saturated, low density, ice matrix-supported crust >1 m deep (see ice axe and footprint for scale). (b) Solid, high-density ice that was protected from shortwave penetration and subsurface melting by a remnant cap of seasonal snow. Photos by (a) Laurence C. Smith and (b) Matthew G. Cooper.

1069

1070   Another, less likely possibility is additional missed leakage in Rio Behar catchment even beyond

1071   the internally drained areas and crevasse fields manually identified and removed from the

1072   WorldView satellite imagery. While crevasse fields were visually excluded from our lower

1073   bound catchment map, some may have been missed, and it is possible that variable water

1074   storage occurs within crevasses which grow and/or fluctuate in volume with changing ice

1075   velocities. Field observations and repeat satellite/UAV imaging of solitary cracks/lineaments

1076   running through our ADCP river reach and elsewhere strongly suggests that these features are

1077   sealed: none developed into moulins and remotely sensed wetted flow widths were

1078   indistinguishable immediately upstream and downstream of these lineaments, indicating no

1079   reductions in discharge.

1080

1081   **Supplementary discussion II: Scale issue and comparison with GRACE studies**

1082   <u>Size of the study area vs. SMB model domain:</u> As noted in the main text, Behar catchment has

1083   an area between 51.4 and 69.1 km^2 (with a best estimate 63.1 km^2). Other IDCs on the GrIS

1084   surface typically have areas of 10s of km^2 (e.g. ranging from 0.4 – 244.9 km^2 for 799 IDCs(7)).

1085   The described procedure for directly measuring runoff R integrates over these scales. As such,

1086   the measurements, coefficients, and fluvial processes described in this paper are broader in

1087   scale than a point-based Automated Weather Station (AWS) and are intrinsic to the horizontal

1088   scale of a single climate model grid cell or perhaps several grid cells. While a significant

1089   improvement over point measurements, this scale is still small relative to the spatial domain

1090   over which climate/SMB models are typically run. This cautions the extent to which conclusions

1091   and interpretations drawn from a small subarea may be generalized. For example, local model

1092   outputs may differ from observations by tens of percent, but increasing the domain area

1093   improves model results (79). Also, there are known local variations in model performance, for

1094   example at the southern tip of the GrIS, melt and runoff are likely both underestimated due to

1095   sensible heat flux underestimation (80). The RACMO2.3 1-km downscaled product suggests

1096   underestimation in melt and runoff at the ice sheet margins for the original output at 11-km

1097   resolution owing to too low bare ice albedo and a relatively coarse topography in the low

1098   ablation zone. Viewed from this perspective, even the quite large (from an *in situ* perspective)

1099   Rio Behar catchment is too small to make generalized statements about model performance at

1100   the pan-Greenland scale.

1101   <u>Comparison with GRACE:</u> The second way in which our findings are supported across a broader

1102   geographic scale is through comparison of GRACE (Gravity Recovery and Climate Experiment)

1103   satellite retrievals across melt-intensive sectors of the GrIS ablation zone. In particular, we

1104   examine the important intercomparison studies of GRACE vs. SMB mass loss (or more precisely,

1105   SMB minus ice discharge SMB-D) of Sasgen *et al.* (81) and Xu *et al.* (82). Behar catchment and

1106   its surrounding IDCs (covering of 13,563 km^2, see above) are contained within Sector F of

1107   Sasgen *et al.* (81) (covering 417,000 km^2, the second-largest sector in the study) and Sector DS6

1108   of Xu *et al.* (82) (area not provided but similar to Sasgen *et al.* (81)). From Sasgen *et al.* (81),

1109   climate model predictions significantly overestimate mass loss relative to GRACE in this melt-

1110   intensive sector (-66 Gt/yr SMB-D vs. -45 $\pm$8 Gt/yr for GRACE, Table 2). A similar result is found

1111   by Xu *et al.* (82) using the input-output method (IOM), which quantifies the difference between

1112   mass input and output by studying SMB-D. They find IOM mass losses of -14 $\pm$8 Gt/yr (vs. -6 $\pm$9

1113   Gt/yr for GRACE), -32 $\pm$12 Gt/yr (vs. -24 $\pm$8 Gt/yr) and -46 $\pm$14 Gt/yr (vs. -38 $\pm$8 Gt/yr) for time

1114   periods 2003-2007, 2003-2010, and 2003-2012, respectively (Xu *et al.*, Table 1, Sector DS6) (82).

1115   In all cases, surface mass balance over-predicts mass loss relative to GRACE in

1116   western/southwestern Greenland; it is only through model under-prediction of mass loss in

1117   other sectors (notably Sector G in Sasgen et al., and Sector DS8 in Xu *et al.* (82), both northwest

1118   Greenland) that these large differences cancel fortuitously, lending the conclusion of

1119   Greenland-wide agreement between GRACE and SMB-D models. Despite the large

1120   uncertainties associated with these estimates and the much larger geographic areas (mascons)

1121   studied by GRACE, we determine from these studies that SMB-modeled mass losses exceed

1122   GRACE derived mass losses over melt-intensive west/southwest Greenland, consistent with the

1123   findings of this study.

1124

1125   **Supplementary Figures (see following pages)**

1126
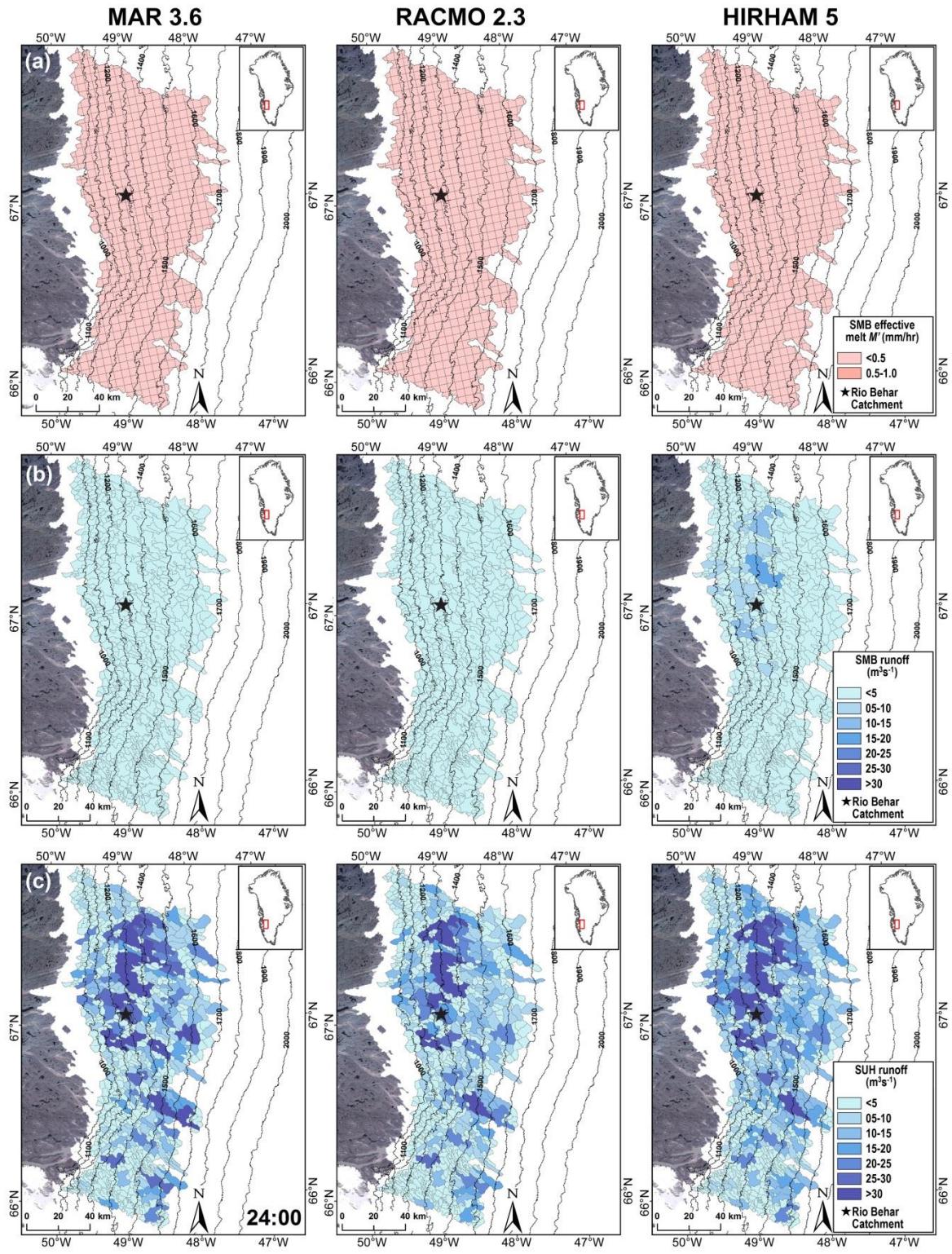
1127

Figure S11: Companion to Figure 5, showing nighttime runoff ten hours later (02:00 on 22 July 2015) for (a) MAR3.6, RACMO2.3, and HIRHAM5 climate/SMB model outputs of corrected meltwater production (*M'*); (b) instantaneous area-integrated runoff; and (c) SUH-

1131 **attenuated runoff. Shutdowns of (a) melt production, and (b) runoff are not present in (c),**
1132 **owing to continued supraglacial river discharge entering moulins at night. MERRA-2 is not**
1133 **shown because it does not output *M*; point SEB is not shown because its output is not gridded. Note**
1134 **that output (a) has units of water depth equivalent (mm hr^{-1}) and units of discharge (m^3 s^{-1})**
1135 **following intersection with supraglacial catchment boundaries (b), (c).   Black star at**
1136 **approximately 67N, 49W denotes Rio Behar catchment.**

1137
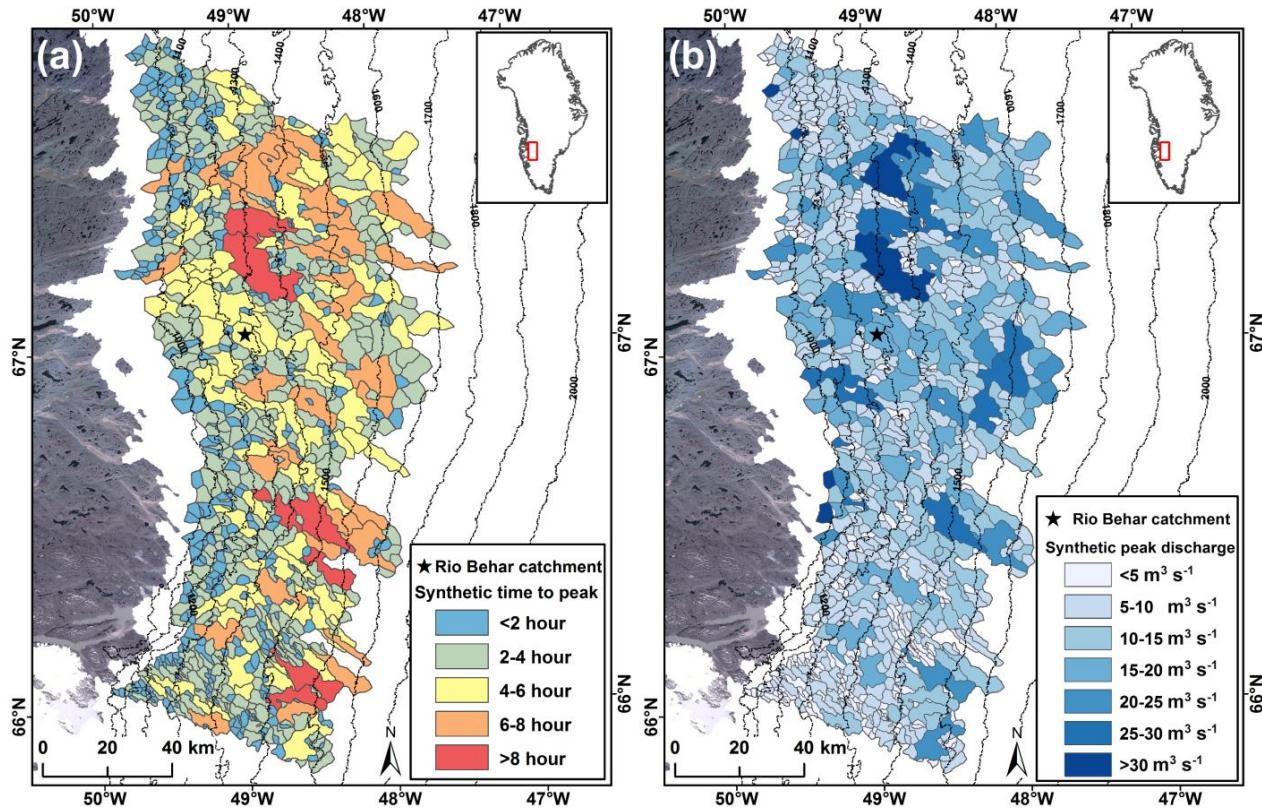
1138

1139 **Figure S12:   Recolored version of Figure 4 for readers having color vision deficiency (for**
1140 **caption see Figure 4 main text)**

1141

1142 **Supplementary Notes/Acknowledgments**

1157

1158

Supplementary References

1160   1.   SonTek (2016) RiverSurveyor S5/M9 System Manual Firmware Version 3.96.
1161   2.   Braithwaite RJ, Konzelmann T, Marty C, & Olesen OB (1998) Errors in daily ablation measurements in northern Greenland, 1993-94, and their implications for glacier climate stuthes. *J. Glaciol.* 44(148):583-588.
1164   3.   Cooper MG*, et al.* (2017) Near-surface meltwater storage in low density bare ice of the Greenland ice sheet ablation zone. *Cryosph. Discuss.*
1166   4.   Ryan JC*, et al.* (2015) UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet. *Cryosph.* 9(1):1-11.
1168   5.   Chandler DM*, et al.* (2013) Evolution of the subglacial drainage system beneath the Greenland Ice Sheet revealed by tracers. *Nature Geosci.* 6(3):195-198.
1170   6.   McGrath D, Colgan W, Steffen K, Lauffenburger P, & Balog J (2011) Assessing the summer water budget of a moulin basin in the Sermeq Avannarleq ablation region, Greenland ice sheet. *J. Glaciol.* 57(205):954-964.
1173   7.   Yang K & Smith LC (2016) Internally drained catchments dominate supraglacial hydrology of the southwest Greenland Ice Sheet. *J. Geophys. Res. Earth Surf.* 121:1891-1910.
1175   8.   AgiSoft (2013) AgiSoft PhotoScan.
1176   9.   Shean DE*, et al.* (2016) An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. *ISPRS J. Photogramm. Remote Sens.* 116:101-117.
1179   10.   Willis MJ, Herried BG, Bevis MG, & Bell RE (2015) Recharge of a subglacial lake by surface meltwater in northeast Greenland. *Nature* 518(7538):223-227.
1181   11.   Karlstrom L & Yang K (2016) Fluvial supraglacial landscape evolution on the Greenland Ice Sheet. *Geophys. Res. Lett.* 43:2683–2692.
1183   12.   Yang K, Karlstrom L, Smith LC, & Li M (2017) Automated high resolution satellite image registration using supraglacial rivers on the Greenland Ice Sheet. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 10(3):845-856.
1186   13.   Smith LC*, et al.* (2015) Efficient meltwater drainage through supraglacial streams and rivers on the southwest Greenland ice sheet. *Proc. Natl. Acad. Sci. U.S.A.* 112(4):1001-1006.
1188   14.   Pope A*, et al.* (2016) Estimating supraglacial lake depth in West Greenland using Landsat 8 and comparison with other multispectral methods. *Cryosph.* 10(1):15-27.
1190   15.   Pedregosa F*, et al.* (2011) Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12:2825-2830.
1192   16.   Chow VT (1964) *Handbook of Applied Hydrology: A Compendium of Water-resources Technology* (McGraw-Hill Company) 1st Ed.
1194   17.   Yates P & Snyder WM (1975) Predicting recessions through convolution. *Water Resour. Res.* 11(3):418-422.
1196   18.   Tallaksen LM (1995) A review of baseflow recession analysis. *J. Hydrol.* 165(1):349-370.
1197   19.   Dingman SL (2015) *Physical hydrology (3rd edition)* (Waveland press).
1198   20.   Collins WT (1939) Runoff distribution graphs from precipitation occurring in more than one time unit. *Civil Eng.* 9(9):559-561.
1200   21.   Snyder FF (1938) Synthetic unit-graphs. *Eos, Trans. Amer. Geophys. Union* 19(1):447-454.
1201   22.   Singh PK, Mishra SK, & Jain MK (2014) A review of the synthetic unit hydrograph: from the empirical UH to advanced geomorphological methods. *Hydrolog. Sci. J.* 59(2):239-261.
1203   23.   Yang K, Smith LC, Chu VW, Gleason CJ, & Li M (2015) A Caution on the Use of Surface Digital Elevation Models to Simulate Supraglacial Hydrology of the Greenland Ice Sheet. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 8(11):5212-5224.

1206   24.   Banwell AF, Arnold NS, Willis IC, Tedesco M, & Ahlstrøm AP (2012) Modeling supraglacial water
1207   routing and lake filling on the Greenland Ice Sheet. *J. Geophys. Res.* 117(F4):F04012.

1208   25.   Linsley RK (1943) Application of the synthetic unit-graph in the western mountain states. *Eos,*
1209   *Trans. Amer. Geophys. Union* 24(2):580-587.

1210   26.   Cordery I (1968) Synthetic unit hydrographs for small catchments in Eastern New South Wales.
1211   *Civil Engineering Transactions, Institution of Engineers (Australia)* 10:47-58.

1212   27.   Wyatt FR & Sharp MJ (2015) Linking surface hydrology to flow regimes and patterns of velocity
1213   variability on Devon Ice Cap, Nunavut. *J. Glaciol.* 61(226):387.

1214   28.   Andrews LC*, et al.* (2014) Direct observations of evolving subglacial drainage beneath the
1215   Greenland Ice Sheet. *Nature* 514(7520):80-83.

1216   29.   van den Broeke M*, et al.* (2008) Partitioning of melt energy and meltwater fluxes in the ablation
1217   zone of the west Greenland ice sheet. *Cryosph.* 2(2):179-189.

1218   30.   Christensen OB*, et al.* (2006) The HIRHAM5 Regional Climate Model. Version 5. in *Danish*
1219   *Meteorological Institute Technical Report*.

1220   31.   Lucas-Picher P*, et al.* (2012) Very high resolution regional climate model simulations over
1221   Greenland: Identifying added value. *J. Geophys. Res. Atmos.* 117:D02108.

1222   32.   Undén P*, et al.* (2002) HIRLAM-5 Scientific Documentation. Scientific Report.

1223   33.   Undèn P*, et al.* (2002) HIRLAM-5 Scientific Documentation. Technical Report.

1224   34.   Roeckner E*, et al.* (2003) The atmospheric general circulation model ECHAM 5. PART I: Model
1225   description. (Max-Planck-Institut für Meteorologie (MPI-M)).

1226   35.   Dee DP*, et al.* (2011) The ERA-Interim reanalysis: configuration and performance of the data
1227   assimilation system. *Q. J. R. Meteorol. Soc.* 137(656):553-597.

1228   36.   Langen PL*, et al.* (2015) Quantifying Energy and Mass Fluxes Controlling Godthåbsfjord
1229   Freshwater Input in a 5-km Simulation (1991–2012). *J. Clim.* 28(9):3694-3713.

1230   37.   Box JE*, et al.* (2012) Greenland ice sheet albedo feedback: thermodynamics and atmospheric
1231   drivers. *Cryosph.* 6(4):821-839.

1232   38.   Vionnet V*, et al.* (2012) The detailed snowpack scheme Crocus and its implementation in SURFEX
1233   v7.2. *Geosci. Model Dev.* 5(3):773-791.

1234   39.   Lefebre F, Gallée H, van Ypersele J-P, & Greuell W (2003) Modeling of snow and ice melt at ETH
1235   Camp (West Greenland): A study of surface albedo. *J. Geophys. Res. Atmos.* 108(D8):4231.

1236   40.   Zuo Z & Oerlemans J (1996) Modelling albedo and specific balance of the Greenland ice sheet:
1237   calculations for the Sendre Stromøord transect. *J. Glaciol.* 42(141):305-317.

1238   41.   Fettweis X, Gallée H, Lefebre F, & Ypersele J-P (2005) Greenland surface mass balance simulated
1239   by a regional climate model and comparison with satellite-derived data in 1990–1991. *Climate*
1240   *Dynamics* 24(6):623-640.

1241   42.   Fettweis X*, et al.* (2013) Estimating the Greenland ice sheet surface mass balance contribution to
1242   future sea level rise using the regional atmospheric climate model MAR. *Cryosph.* 7(2):469-489.

1243   43.   Ridder KD & Gallée H (1998) Land Surface–Induced Regional Climate Change in Southern Israel. *J.*
1244   *Appl. Meteorol.* 37(11):1470-1485.

1245   44.   Gallée H & Schayes G (1994) Development of a three-dimensional meso-γ primitive equation
1246   model: katabatic winds simulation in the area of Terra Nova Bay, Antarctica. *Mon. Weather Rev.*
1247   122(4):671-685.

1248   45.   Brun E, David P, Sudul M, & Brunot G (1992) A numerical model to simulate snow-cover
1249   stratigraphy for operational avalanche forecasting. *J. Glaciol.* 38(128):13-22.

1250   46.   Gallée H, Guyomarc'h G, & Brun E (2001) Impact Of Snow Drift On The Antarctic Ice Sheet
1251   Surface Mass Balance: Possible Sensitivity To Snow-Surface Properties. *Bound.-Layer Meteor.*
1252   99(1):1-19.

1253   47.   Fettweis X*, et al.* (2017) Reconstructions of the 1900–2015 Greenland ice sheet surface mass
1254   balance using the regional climate MAR model. *Cryosph.* 11(2):1015-1033.
1255   48.   Bamber JL, Ekholm S, & Krabill WB (2001) A new, high-resolution digital elevation model of
1256   Greenland fully validated with airborne laser altimeter data. *J. Geophys. Res. Solid Earth*
1257   106(B4):6733-6745.
1258   49.   Bamber JL, Layberry RL, & Gogineni SP (2001) A new ice thickness and bed data set for the
1259   Greenland ice sheet 1. Measurement, data reduction, and errors. *J. Geophys. Res. Atmos.*
1260   106(D24):33773-33780.
1261   50.   Van Meijgaard E*, et al.* (2008) Technical Report: The KNMI regional atmospheric climate model
1262   RACMO version 2.1. (Royal Netherlands Meteorological Institute, De Bilt).
1263   51.   Ettema J, van den Broeke MR, van Meijgaard E, & van de Berg WJ (2010) Climate of the
1264   Greenland ice sheet using a high-resolution climate model – Part 2: Near-surface climate and
1265   energy balance. *Cryosph.* 4(4):529-544.
1266   52.   Kuipers Munneke P*, et al.* (2011) A new albedo parameterization for use in climate models over
1267   the Antarctic ice sheet. *J. Geophys. Res. Atmos.* 116:D05114.
1268   53.   Kuipers Munneke P*, et al.* (2015) Elevation change of the Greenland Ice Sheet due to surface
1269   mass balance and firn processes, 1960-2014. *Cryosph.* 9(6):2009-2025.
1270   54.   Lenaerts JTM, van den Broeke MR, van Angelen JH, van Meijgaard E, & Déry SJ (2012) Drifting
1271   snow climate of the Greenland ice sheet: a study with a regional climate model. *Cryosph.*
1272   6(4):891-899.
1273   55.   Noël B*, et al.* (2016) A daily, 1-km resolution dataset of downscaled Greenland ice sheet surface
1274   mass balance (1958-2015). *Cryosph.* 2016:1-29.
1275   56.   Van Wessem J*, et al.* (2014) Improved representation of East Antarctic surface mass balance in a
1276   regional atmospheric climate model. *J. Glaciol.* 60(222):761-770.
1277   57.   Forster RR*, et al.* (2014) Extensive liquid meltwater storage in firn within the Greenland ice sheet.
1278   *Nature Geosci.* 7(2):95-98.
1279   58.   Bosilovich MG*, et al.* (2016) MERRA-2. Initial evaluation of the climate. in *Technical Report Series*
1280   *on Global Modeling and Data Assimilation*, ed Koster RD (National Aeronautics and Space
1281   Administration, Greenbelt, Maryland).
1282   59.   Rienecker MM*, et al.* (2011) MERRA: NASA's modern-era retrospective analysis for research and
1283   applications. *J. Clim.* 24(14):3624-3648.
1284   60.   Lin S-J (2004) A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon.*
1285   *Weather Rev.* 132(10):2293-2307.
1286   61.   Molod A, Takacs L, Suarez M, & Bacmeister J (2015) Development of the GEOS-5 atmospheric
1287   general circulation model: evolution from MERRA to MERRA2. *Geosci. Model Dev.* 8(5):1339-
1288   1356.
1289   62.   Takacs LL, Suárez MJ, & Todling R (2016) Maintaining atmospheric mass and water balance in
1290   reanalyses. *Q. J. R. Meteorol. Soc.* 142(697):1565-1573.
1291   63.   Cullather RI, Nowicki SMJ, Zhao B, & Suarez MJ (2014) Evaluation of the Surface Representation
1292   of the Greenland Ice Sheet in a General Circulation Model. *J. Clim.* 27(13):4835-4856.
1293   64.   Stieglitz M, Ducharne A, Koster R, & Suarez M (2001) The impact of detailed snow physics on the
1294   simulation of snow cover and subsurface thermodynamics at continental scales. *J.*
1295   *Hydrometeorol.* 2(3):228-242.
1296   65.   Lynch-Stieglitz M (1994) The Development and Validation of a Simple Snow Model for the GISS
1297   GCM. *J. Clim.* 7(12):1842-1855.
1298   66.   Greuell W & Konzelmann T (1994) Greenland ice margin experiment (GIMEx) Numerical
1299   modelling of the energy balance and the englacial temperature of the Greenland Ice Sheet.

1300   Calculations for the ETH-Camp location (West Greenland, 1155 m a.s.l.). *Glob. Planet. Change*
1301   9(1):91-114.
1302 67. Jordan R (1991) A one-dimensional temperature model for a snow cover: Technical
1303   documentation for SNTHERM.89. (U.S. Army Corps of Engineers, Cold Regions Research and
1304   Engineering Laboratory).
1305 68. Van As D (2011) Warming, glacier melt and surface energy budget from weather station
1306   observations in the Melville Bay region of northwest Greenland. *J. Glaciol.* 57(202):208-220.
1307 69. Smeets CJPP & van den Broeke MR (2008) Temporal and Spatial Variations of the Aerodynamic
1308   Roughness Length in the Ablation Zone of the Greenland Ice Sheet. *Bound.-Layer Meteor.*
1309   128(3):315-338.
1310 70. Snyder JP (1987) Map projections: A working manual. in *USGS Professional Paper* (Washington,
1311   D.C.).
1312 71. Pitcher LH, Smith LC, & Gleason CJ (2016) CryoSheds: a GIS modeling framework for delineating
1313   land-ice watersheds for the Greenland Ice Sheet. *GIsci. Remote Sens.* 53(6):707-722.
1314 72. Irvine-Fynn TDL, Hodson AJ, Moorman BJ, Vatne G, & Hubbard AL (2011) Polythermal glacier
1315   hydrology: a review. *Rev. Geophys.* 49(4):RG4002.
1316 73. Cook JM, Hodson AJ, & Irvine-Fynn TDL (2016) Supraglacial weathering crust dynamics inferred
1317   from cryoconite hole hydrology. *Hydrol. Process.* 30(3):433-446.
1318 74. Hoffman MJ, Fountain AG, & Liston GE (2014) Near-surface internal melting: a substantial mass
1319   loss on Antarctic Dry Valley glaciers. *J. Glaciol.* 60(220):361-374.
1320 75. Müller F & Keeler C (1969) Errors in Short-Term Ablation Measurements on Melting Ice Surfaces.
1321   *J. Glaciol.* 8(52):91-105.
1322 76. Karlstrom L, Zok A, & Manga M (2014) Near-surface permeability in a supraglacial drainage basin
1323   on the Llewellyn Glacier, Juneau Icefield, British Columbia. *Cryosph.* 8(2):537-546.
1324 77. Fountain AG & Walder JS (1998) Water flow through temperate glaciers. *Rev. Geophys.*
1325   36(3):299-328.
1326 78. Willis IC, Arnold NS, & Brock BW (2002) Effect of snowpack removal on energy balance, melt and
1327   runoff in a small supraglacial catchment. *Hydrol. Process.* 16(14):2721-2749.
1328 79. Van As D*, et al.* (2014) Increasing meltwater discharge from the Nuuk region of the Greenland
1329   ice sheet and implications for mass balance (1960–2012). *J. Glaciol.* 60(220):314-322.
1330 80. Fausto RS*, et al.* (2016) The implication of nonradiative energy fluxes dominating Greenland ice
1331   sheet exceptional ablation area surface melt in 2012. *Geophys. Res. Lett.* 43(6):2649-2658.
1332 81. Sasgen I*, et al.* (2012) Timing and origin of recent regional ice-mass loss in Greenland. *Earth
1333   Planet. Sci. Lett.* 333–334:293-303.
1334 82. Xu Z, Schrama EJO, van der Wal W, van den Broeke M, & Enderlin EM (2016) Improved GRACE
1335   regional mass balance estimates of the Greenland ice sheet cross-validated with the input–
1336   output method. *Cryosph.* 10(2):895-912.

1337

1. Van den Broeke M, *et al.* (2009) Partitioning recent Greenland mass loss. *Science* 326(5955):984-986.

2. Shepherd A, *et al.* (2012) A reconciled estimate of ice-sheet mass balance. *Science* 338(6111):1183-1189.

3. Andersen ML, *et al.* (2015) Basin-scale partitioning of Greenland ice sheet mass balance components (2007–2011). *Earth Planet. Sci. Lett.* 409(0):89-95.

4. Enderlin EM, *et al.* (2014) An improved mass budget for the Greenland Ice Sheet. *Geophys. Res. Lett.* 41(3):866–872.

5. Chu VW (2014) Greenland ice sheet hydrology: a review. *Prog. Phys. Geogr.* 38(1):19-54.

6. Irvine-Fynn TDL, Hodson AJ, Moorman BJ, Vatne G, & Hubbard AL (2011) Polythermal glacier hydrology: a review. *Rev. Geophys.* 49(4):RG4002.

7. Smith LC, *et al.* (2015) Efficient meltwater drainage through supraglacial streams and rivers on the southwest Greenland ice sheet. *Proc. Natl. Acad. Sci. U.S.A.* 112(4):1001-1006.

8. McGrath D, Colgan W, Steffen K, Lauffenburger P, & Balog J (2011) Assessing the summer water budget of a moulin basin in the Sermeq Avannarleq ablation region, Greenland ice sheet. *J. Glaciol.* 57(205):954-964.

9. Zwally HJ, *et al.* (2002) Surface melt-induced acceleration of Greenland ice-sheet flow. *Science* 297(5579):218-222.

10. Bartholomaus TC, Anderson RS, & Anderson SP (2008) Response of glacier basal motion to transient water storage. *Nature Geosci.* 1(1):33-37.

11. Van de Wal RSW, *et al.* (2008) Large and rapid melt-induced velocity changes in the ablation zone of the Greenland ice sheet. *Science* 321(5885):111-113.

12. Schoof C (2010) Ice-sheet acceleration driven by melt supply variability. *Nature* 468(7325):803-806.

13. Kulessa B, *et al.* (2017) Seismic evidence for complex sedimentary control of Greenland Ice Sheet flow. *Sci. Adv.* 3(8):e1603071.

14. Lenaerts JTM, *et al.* (2015) Representing Greenland ice sheet freshwater fluxes in climate models. *Geophys. Res. Lett.* 42(15):6373-6381.

15. Machguth H, *et al.* (2016) Greenland meltwater storage in firn limited by near-surface ice formation. *Nature Clim. Change* 6(4):390-393.

16. Stokes CR, Margold M, Clark CD, & Tarasov L (2016) Ice stream activity scaled to ice sheet volume during Laurentide Ice Sheet deglaciation. *Nature* 530(7590):322-326.

17. Rennermalm AK, *et al.* (2013) Evidence of meltwater retention within the Greenland ice sheet. *Cryosph.* 7(5):1433-1445.

18. Van As D, *et al.* (2014) Increasing meltwater discharge from the Nuuk region of the Greenland ice sheet and implications for mass balance (1960–2012). *J. Glaciol.* 60(220):314-322.

19. Overeem I, *et al.* (2015) River inundation suggests ice-sheet runoff retention. *J. Glaciol.* 61(228):776-788.

20. Bartholomew I, *et al.* (2011) Supraglacial forcing of subglacial drainage in the ablation zone of the Greenland ice sheet. *Geophys. Res. Lett.* 38.

21. Catania GA & Neumann TA (2010) Persistent englacial drainage features in the Greenland Ice Sheet. *Geophys. Res. Lett.* 37(2):L02501.

22. Covington MD, *et al.* (2012) Quantifying the effects of glacier conduit geometry and recharge on proglacial hydrograph form. *J. Hydrol.* 414–415:59-71.

23. Wright PJ, Harper JT, Humphrey NF, & Meierbachtol TW (2016) Measured basal water pressure variability of the western Greenland Ice Sheet: Implications for hydraulic potential. *J. Geophys. Res. Earth Surf.* 121(6):1134-1147.

24. Fitzpatrick AAW, *et al.* (2014) A decade (2002–2012) of supraglacial lake volume estimates across Russell Glacier, West Greenland. *Cryosph.* 8:107-121.

25. Banwell AF, Willis IC, & Arnold NS (2013) Modeling subglacial water routing at Paakitsoq, W Greenland. *J. Geophys. Res. Earth Surf.* 118(3):1282-1295.

26. Pitcher LH, Smith LC, & Gleason CJ (2016) CryoSheds: a GIS modeling framework for delineating land-ice watersheds for the Greenland Ice Sheet. *GIsci. Remote Sens.* 53(6):707-722.

27. Lindbäck K, *et al.* (2015) Subglacial water drainage, storage, and piracy beneath the Greenland Ice Sheet. *Geophys. Res. Lett.* 42(18):7606-7614.

28. Chu W, Creyts TT, & Bell RE (2016) Rerouting of subglacial water flow between neighboring glaciers in West Greenland. *J. Geophys. Res. Earth Surf.* 121(5):925-938.

29. Gleason CJ, *et al.* (2015) Technical Note: Semi-automated effective width extraction from time-lapse RGB imagery of a remote, braided Greenlandic river. *Hydrol. Earth Syst. Sci.* 19(6):2963-2969.

30. Hoffman MJ, Catania GA, Neumann TA, Andrews LC, & Rumrill JA (2011) Links between acceleration, melting, and supraglacial lake drainage of the western Greenland Ice Sheet. *J. Geophys. Res. Earth Surf.* 116(F4):F04035.

31. Bartholomew I, *et al.* (2012) Short-term variability in Greenland Ice Sheet motion forced by time-varying meltwater drainage: Implications for the relationship between subglacial drainage system behavior and ice velocity. *J. Geophys. Res. Earth Surf.* 117:F03002.

32. Chandler DM, *et al.* (2013) Evolution of the subglacial drainage system beneath the Greenland Ice Sheet revealed by tracers. *Nature Geosci.* 6(3):195-198.

33. Andrews LC, *et al.* (2014) Direct observations of evolving subglacial drainage beneath the Greenland Ice Sheet. *Nature* 514(7520):80-83.

34. van de Wal RSW, *et al.* (2015) Self-regulation of ice flow varies across the ablation area in south-west Greenland. *Cryosph.* 9(2):603-611.

35. Fausto RS, *et al.* (2016) The implication of nonradiative energy fluxes dominating Greenland ice sheet exceptional ablation area surface melt in 2012. *Geophys. Res. Lett.* 43(6):2649-2658.

36. van den Broeke MR, Smeets CJPP, & van de Wal RSW (2011) The seasonal cycle and interannual variability of surface energy balance and melt in the ablation zone of the west Greenland ice sheet. *Cryosph.* 5(2):377-390.

37. Nghiem SV, Steffen K, Kwok R, & Tsai WY (2001) Detection of snowmelt regions on the Greenland ice sheet using diurnal backscatter change. *J. Glaciol.* 47(159):539-547.

38. Hubbard BP, Sharp MJ, Willis IC, Nielsen MK, & Smart CC (1995) Borehole water-level variations and the structure of the subglacial hydrological system of Haut Glacier d'Arolla, Valais, Switzerland. *J. Glaciol.* 41(139):572-583.

39. Werder MA, Hewitt IJ, Schoof CG, & Flowers GE (2013) Modeling channelized and distributed subglacial drainage in two dimensions. *J. Geophys. Res. Earth Surf.* 118(4):2140-2158.

40. Hewitt IJ (2013) Seasonal changes in ice sheet motion due to melt water lubrication. *Earth Planet. Sci. Lett.* 371–372:16-25.

41. van As D, *et al.* (2017) Hypsometric amplification and routing moderation of Greenland ice sheet meltwater release. *Cryosph.* 11(3):1371-1386.

42. Steger CR, *et al.* (2017) Firn Meltwater Retention on the Greenland Ice Sheet: A Model Comparison. *Front. Earth Sci.* 5:16.

43. Yang K & Smith LC (2016) Internally drained catchments dominate supraglacial hydrology of the southwest Greenland Ice Sheet. *J. Geophys. Res. Earth Surf.* 121:1891-1910.

44. Snyder FF (1938) Synthetic unit-graphs. *Eos, Trans. Amer. Geophys. Union* 19(1):447-454.

45. Chow VT (1964) *Handbook of Applied Hydrology: A Compendium of Water-resources Technology* (McGraw-Hill Company) 1st Ed.

46. Yang K, Smith LC, Chu VW, Gleason CJ, & Li M (2015) A Caution on the Use of Surface Digital Elevation Models to Simulate Supraglacial Hydrology of the Greenland Ice Sheet. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 8(11):5212-5224.

47. Singh PK, Mishra SK, & Jain MK (2014) A review of the synthetic unit hydrograph: from the empirical UH to advanced geomorphological methods. *Hydrolog. Sci. J.* 59(2):239-261.

48. Poinar K, *et al.* (2015) Limits to future expansion of surface-melt-enhanced ice flow into the interior of western Greenland. *Geophys. Res. Lett.* 42:1800-1807.

49. Wyatt FR & Sharp MJ (2015) Linking surface hydrology to flow regimes and patterns of velocity variability on Devon Ice Cap, Nunavut. *J. Glaciol.* 61(226):387.

50. Cooper MG, *et al.* (2017) Near-surface meltwater storage in low density bare ice of the Greenland ice sheet ablation zone. *Cryosph. Discuss.*

51. Braithwaite RJ, Konzelmann T, Marty C, & Olesen OB (1998) Errors in daily ablation measurements in northern Greenland, 1993-94, and their implications for glacier climate stuthes. *J. Glaciol.* 44(148):583-588.

52. Bennartz R, *et al.* (2013) July 2012 Greenland melt extent enhanced by low-level liquid clouds. *Nature* 496(7443):83-86.

53. van den Broeke M, *et al.* (2008) Partitioning of melt energy and meltwater fluxes in the ablation zone of the west Greenland ice sheet. *Cryosph.* 2(2):179-189.

54. Cook JM, Hodson AJ, & Irvine-Fynn TDL (2016) Supraglacial weathering crust dynamics inferred from cryoconite hole hydrology. *Hydrol. Process.* 30(3):433-446.

55. Willis IC, Arnold NS, & Brock BW (2002) Effect of snowpack removal on energy balance, melt and runoff in a small supraglacial catchment. *Hydrol. Process.* 16(14):2721-2749.

56. Shepherd A, *et al.* (2009) Greenland ice sheet motion coupled with daily melting in late summer. *Geophys. Res. Lett.* 36(1):L01501.

57. Gleason CJ, *et al.* (2016) Characterizing supraglacial meltwater channel hydraulics on the Greenland Ice Sheet from in situ observations. *Earth Surf. Process. Landf.*:n/a-n/a.

58. Leeson AA, Shepherd A, Palmer S, Sundal A, & Fettweis X (2012) Simulating the growth of supraglacial lakes at the western margin of the Greenland ice sheet. *Cryosph.* 6(5):1077-1086.

59. Tedesco M, *et al.* (2013) Evidence and analysis of 2012 Greenland records from spaceborne observations, a regional climate model and reanalysis data. *Cryosph.* 7(2):615-630.

60. Hoffman MJ, Fountain AG, & Liston GE (2014) Near-surface internal melting: a substantial mass loss on Antarctic Dry Valley glaciers. *J. Glaciol.* 60(220):361-374.

61. Müller F & Keeler C (1969) Errors in Short-Term Ablation Measurements on Melting Ice Surfaces. *J. Glaciol.* 8(52):91-105.

62. Sasgen I, *et al.* (2012) Timing and origin of recent regional ice-mass loss in Greenland. *Earth Planet. Sci. Lett.* 333–334:293-303.

63. Xu Z, Schrama EJO, van der Wal W, van den Broeke M, & Enderlin EM (2016) Improved GRACE regional mass balance estimates of the Greenland ice sheet cross-validated with the input–output method. *Cryosph.* 10(2):895-912.

64. Vernon CL, *et al.* (2013) Surface mass balance model intercomparison for the Greenland ice sheet. *Cryosph.* 7(2):599-614.